

UNIVERSIDADE DE LISBOA
Faculdade de Ciências
Departamento de Informática



Metadata – Reporting de Gestão

Lídia Simões Raimundo

Mestrado em Engenharia Informática

2007

UNIVERSIDADE DE LISBOA
Faculdade de Ciências
Departamento de Informática



Metadata – Reporting de Gestão

Lídia Simões Raimundo

**Projecto orientado pelo Prof. Dr. Francisco José Moreira Couto
e co-orientado pela Dra. Luísa Maria de Brito da Costa Ferreira**

Mestrado em Engenharia Informática
2007

Declaração

Lídia Simões Raimundo, aluno nº 27780 da Faculdade de Ciências da Universidade de Lisboa, declara ceder os seus direitos de cópia sobre o seu Relatório de Projecto em Engenharia Informática, intitulado "Metadata – Reporting de Gestão" (versão pública), realizado no ano lectivo de 2006/2007 à Faculdade de Ciências da Universidade de Lisboa para o efeito de arquivo e consulta nas suas bibliotecas e publicação do mesmo em formato electrónico na Internet.

FCUL, 22 de Outubro de 2007.

Dra. Luísa Maria de Brito da Costa Ferreira, supervisor do projecto de Lídia Simões Raimundo, aluna da Faculdade de Ciências da Universidade de Lisboa, declara concordar com a divulgação do Relatório do Projecto em Engenharia Informática, intitulado "Metadata – Reporting de Gestão" (versão pública).

Lisboa, 22 de Outubro de 2007.

*"The artist is nothing without the gift,
but the gift is nothing without work."*

- Emile Zola

Agradecimentos

Gostaria de agradecer a todas as pessoas que contribuíram para a realização deste trabalho.

Agradeço ao meu orientador, Francisco Couto, pela paciência, disponibilidade e apoio sempre presentes. À minha supervisora da instituição de acolhimento, Maria Guerra, o meu obrigada pela sua perseverança, dedicação e críticas construtivas, que contribuíram, em muito, para o meu crescimento a nível profissional e pessoal. À orientadora Luísa Ferreira, agradeço a sua disponibilidade, atenção e constante motivação. Ao meu colega João Candeias agradeço toda a orientação e auxílio prestado numa fase inicial de integração no ambiente de trabalho e no decurso de todo o projecto.

Agradeço a todos os meus colegas por tornarem possível a aquisição de uma postura de trabalho ímpar, que me prepara para todo o trabalho que desenvolverei de futuro.

A todas as pessoas, para mim especiais, a minha maior gratidão, o vosso apoio e motivação reflectem-se neste trabalho.

Resumo

Esta tese descreve o projecto realizado no âmbito da disciplina Projecto em Engenharia Informática do Mestrado em Engenharia Informática da Faculdade de Ciências da Universidade de Lisboa.

Desde o início da década de 90 que as tecnologias Data Warehouse estão amplamente disseminadas, ao nível das organizações. Estes sistemas emergiram, como solução, para sustentar estratégias de mercado e auxiliar o processo de suporte a decisões de uma área de negócios.

O conceito de *Reporting*, tema fulcral neste projecto, traduz-se na restituição de informação aos clientes através da construção e entrega de relatórios com toda a informação por estes requisitada.

Esta dissertação aborda o problema do crescimento exponencial da informação armazenada que é continuamente alterada, num ambiente Data Warehouse. Este processo de contínuo refrescamento da informação, pode colocar em risco o desempenho do sistema e a eficiência do *Reporting*.

Para atenuar este problema, carece-se de um processo complexo de análise e gestão da informação.

Neste trabalho, é proposto um conjunto de técnicas aplicadas à informação existente para que o desempenho do sistema e a eficiência do *Reporting* não fiquem comprometidos pela acumulação de informação obsoleta.

Foi aplicada uma metodologia que combina técnicas de análise e gestão do conhecimento, exploração de informação e manutenção correctiva. Esta metodologia, visa dar resposta ao problema existente, num caso típico de Engenharia do Software.

A viabilidade destas estratégias foi avaliada, quando aplicada a um caso de estudo bem definido, no sentido de verificar a precisão e abrangência dos resultados obtidos.

O projecto descrito nesta tese foi concretizado no âmbito de uma instituição bancária e, consequentemente, a informação manipulada pelo sistema é de natureza financeira.

O trabalho realizado, visa atingir um ambiente organizado, eficiente, reutilizável, mais robusto e caracterizado pela qualidade, contemplando também a redução do espaço físico, ocupado em memória, como um requisito aplicacional a alcançar.

Palavras-chave: Data Warehouse, *Reporting*, Business Intelligence, Análise e Gestão de Informação, MicroStrategy.

Abstract

This thesis describes the project carried out in the scope of the discipline Project in Computer Science Engineering of the Master in Computer Science Engineering from the Faculty of Sciences – University of Lisbon.

Since the early nineties, the Data Warehouse Technologies are widely disseminated among the organizations. These systems have emerged, as a solution, to sustain the market strategies and support the business' decisions.

The concept of Reporting, fundamental in this work, translates itself in the information retrieval to the clients through the construction and delivery of reports with all the information required by these clients.

This dissertation approaches the problem of the exponential growth of the information stored and continuously altered, in a Data Warehouse environment. In this way, the performance of the system and the efficiency of the Reporting are can be at risk.

With the purpose of solving this problem, the analysis and management of the information are essential.

In this work it is proposed a set of techniques applied to the existing information, so that the performance of the system and the efficiency of the Reporting are not compromised by the existence of obsolete information.

It was applied a methodology that combines techniques of analysis and knowledge management, exploration of information and corrective maintenance. This methodology has the objective of responding to the existing problem, in a typical case of Software Engineering.

The feasibility of these strategies has been evaluated, when applied to a well defined study case, in order to verify the precision and recall of the results obtained.

The project described in this thesis was developed in the scope of a banking organization and, consequently, the information manipulated by the system is of financial nature.

This work has the purpose of obtaining an environment characterized by its organization, efficiency, reusability, robustness and quality. The reduction of the physical space, occupied in memory, is also an applicational requirement to accomplish.

Key-Words: Data Warehouse, Reporting, Business Intelligence, Information Analysis and Management, MicroStrategy.

Índice

INTRODUÇÃO	21
1.1. Projecto em Engenharia Informática	21
1.2. Motivação	21
1.3. Enquadramento, Contribuições e Resultados	22
1.4. Enquadramento Institucional	23
1.4.1. Detalhes da Estrutura Organizacional.....	24
1.4.2. A Integração na Instituição de Acolhimento Externa.....	25
1.4.3. Competências e Responsabilidades de um Colaborador.....	25
1.5. Estrutura da Dissertação	26
PANORÂMICA	27
2.1. Panorâmica Geral.....	27
2.2. O Conceito Data Warehouse	28
2.2.1. Perspectiva Arquitectural	28
2.2.1.1. Modelação Dimensional	32
2.2.2. Perspectiva Metodológica de Negócio.....	34
2.2. O Conceito de Reporting.....	35
2.3. Reporting – Evolução histórica	35
OBJECTIVOS, METODOLOGIA E PLANEAMENTO	37
3.1. Objectivos.....	37
3.1.1. Solução Proposta em Diversos Níveis	38
3.1.1.1. Arquitectural	39
3.1.1.2. Processual.....	40
3.1.1.3. Conceptual.....	41
3.1.1.4. Aplicacional	41
3.2. Metodologia	42
3.2.1. Metodologia do ciclo de vida do Projecto	43
3.2.2. Metodologia Adoptada – Detalhe	45
3.3. Planeamento.....	47
3.3.1. Plano de Tarefas Inicial	47
3.3.1.1. Plano de Tarefas Efectuado	48
3.3.1.2. Plano de Tarefas Inicial versus Concretizado.....	48
TRABALHO REALIZADO.....	50
4.1. Responsabilidades no Projecto.....	51
4.2. Ferramentas Utilizadas	52
4.2.1. MicroStrategy	52

4.2.2.	TOAD	56
4.2.3.	Folhas de Cálculo.....	56
4.2.4.	Processamento de Texto.....	57
4.2.5.	Visio	57
4.2.6.	Microsoft Project	57
4.3.	Detalhe das Tarefas do Trabalho Realizado	58
4.3.1.	Análise, Especificação de Requisitos e Modelação	58
4.3.2.	Desenvolvimento/Implementação.....	61
4.3.2.1.	SQL: Structured Query Language	62
4.3.2.2.	PL/SQL: Procedural Language/Structured Query Language	63
4.3.2.3.	Análise e Levantamento de Informação	63
4.3.2.4.	Produção de Documentação.....	63
4.3.2.5.	Metodologia Modular	63
4.3.3.	Passagem a Qualificação	64
4.3.4.	Testes.....	64
4.3.5.	Entrada em Produção e Manutenção.....	65
4.4.	Detalhe dos Resultados Obtidos	65
4.4.1.	Resultados Qualitativos	65
4.4.2.	Resultados Quantitativos.....	65
4.4.3.	Memória Física Ganha em Memória.....	68
4.5.	Precisão e Abrangência dos Resultados Obtidos	69
4.5.1.	Cálculos da Precisão e Abrangência:.....	70
4.6.	Principais Contribuições.....	74
4.6.	Análise Crítica	74
4.6.1.	Ferramentas e Tecnologias.....	75
4.6.2.	Trabalho Realizado e Resultados Obtidos	76
	CONCLUSÃO E TRABALHO FUTURO	78
5.1.	Conclusão	78
5.2.	Trabalho Futuro	79
	REFERÊNCIAS.....	80
	ANEXOS	81
	Anexo 1 – TOAD	81
	Anexo 1.1 - Toad: Editor PL/SQL.....	81
	Anexo 1.2 - Toad: Componente de Navegação	81
	Anexo 2 – Arquitetura Física	82
	Anexo 3 – Planeamento: Mapa de Gantt Detalhado	83

Índice de Figuras

Figura 1 – Estrutura Organizacional da Área Informática do Grupo Bancário ..	24
Figura 2 – O Departamento da Instituição Externa: Principais Competências..	24
Figura 3 – Arquitectura da Informação do Departamento.....	25
Figura 4 – Arquitectura Data Warehouse de referência	31
Figura 5 – Esquema em Estrela.....	33
Figura 6 – Esquema em Floco de Neve.....	33
Figura 7 – Necessidades existentes em diferentes níveis	37
Figura 8 – Níveis da Solução proposta.....	38
Figura 9 – Arquitectura Associada ao Projecto	39
Figura 10 – Relação Aplicação – Utilizador Final/Cliente	42
Figura 11 – Fases da Metodologia de Projecto usada no Departamento	43
Figura 12 – Detalhe das Fases da Metodologia de Projecto usada no Departamento	43
Figura 13 – Mapa Gantt do Planeamento inicialmente definido	48
Figura 14 – Resumo de Tarefas e Ferramentas.....	51
Figura 15 – Ambiente MicroStrategy Desktop (*).....	54
Figura 16 – Ambiente de Criação de Relatórios	55
Figura 17 – Ambiente de Desenvolvimento TOAD	56
Figura 18 – Desenho preliminar decorrente da fase de análise.....	59
Figura 19 - Modelo Arquitectural DW.....	59
Figura 20 – Gráfico de Precisão e Abrangência dos Resultados obtidos.....	71
Figura 21 – Gráfico de Precisão e Abrangência dos Resultados obtidos fórmulas adaptadas.....	73

Índice de Tabelas

Tabela 1 – Responsabilidades e Competências de um Colaborador.....	26
Tabela 2 – Componentes do Modelo Dimensional	32
Tabela 3 – Perspectiva Histórica.....	36
Tabela 4 – Arquitectura Existente e Funcionalidades.....	40
Tabela 5 – Definição de cada uma das Fases da Metodologia	44
Tabela 6 – Planeamento Inicial de Tarefas Segmentado em Quatro Fases Principais.....	47
Tabela 7 – Definição de Conceitos MicroStrategy.....	53
Tabela 8 – Regras definidas para cada tipo de Objecto.....	61
Tabela 9 – Nº Total de Objectos existentes no Repositório de Metadados	65
Tabela 10 – Nº de Objectos indicados para Remoção do Repositório de Metadados	66
Tabela 11 – Relatórios Existentes no Repositório de Metadados.....	67
Tabela 12 – Requisito Aplicacional de Diminuição da Carga de Armazenamento	68
Tabela 13 – Resumo de Resultado por Projecto	70
Tabela 14 – Aplicação directa das fórmulas de Precisão e Abrangência.....	71
Tabela 15 – Aplicação das fórmulas de Precisão e Abrangência com ajustamento	72
Tabela 16 – Ferramentas e Tecnologias – Análise Crítica	76

Glossário

Terminologia e Definições

Atributo – consiste na representação e agrupamento de categorias e níveis de dados. Não são numéricos e representam informação por assunto, sob a forma de categorias e níveis para representar e qualificar um grupo de dados. Permitem responder a perguntas sobre um facto. Os atributos contêm os elementos que constam nos relatórios produzidos pelo software de *Reporting*. Os atributos possuem relações que podem ser de 1 para 1, 1 para muitos ou muitos para muitos.

Business Intelligence – conceito associado à gestão de negócio. Consiste num conjunto de sistemas, aplicações e tecnologias usadas com a finalidade de recolher, dar acesso e analisar dados e informações empresariais que se encontram armazenadas numa BD. Esta extracção do conhecimento de negócio produz resultados utilizados pela empresa para facilitar a tomada de decisões.

Cubo – estrutura física, multi-dimensional, de organização e representação de informação a restituir aos utilizadores finais sob a forma de *Reporting*.

Data Mart – modelo de dados orientado a um objectivo específico, de modo a tratar, de forma mais especializada, um conjunto de informação de um Data Warehouse. Arquitecturalmente podem representar-se de diversas formas, residindo em diferentes localizações na arquitectura típica de um Data Warehouse.

Data Warehouse – é um armazém de dados que difere de uma BD tradicional pelas capacidades que se enunciam em seguida.

- Pode ser visto segundo duas perspectivas, como uma metodologia ou como uma arquitectura de informação.
- Integra dados provenientes das mais diversas fontes externas e dos sistemas operacionais, armazena dados históricos e metadados, permite sumariar os dados com vista a obter informações de gestão.

Um Data Warehouse é usado para descrever todo um sistema de *Business Intelligence*, Sistemas de Suporte à Decisão e outros Sistemas de Gestão.

Data Warehousing – o processo de construção e manutenção de Data Warehouses.

Dimensão (ou Tabela Dimensão) – conjunto de atributos que descrevem uma mesma propriedade da ocorrência. Descrevem as condições de negócio e integram o modelo dimensional.

Extract, Transform e Load – vulgarmente conhecido por processo **ETL**. Consiste na extracção de dados de fontes externas, transformação e tratamento dos dados de modo a dar resposta a requisitos de negócio. Este processo pode ser utilizado para popular não só o Data Warehouse como também o ODS e os *Data Marts*.

Facto – conceptualmente pode ser visto como uma medida de registo das ocorrências de negócio. Os factos consistem numa situação única, dada uma certa combinação de atributos. Permitem o acesso a dados armazenados no Data Warehouse e consistem na base dos relatórios e na maioria das análises requeridas pelos utilizadores. Na maioria dos casos são valores numéricos.

Factual (ou Tabela de Factos) – tabela do modelo dimensional do Data Warehouse físico, tipicamente o esquema em de estrela ou floco-de-neve. Destinam-se a armazenar os diferentes atributos sobre factos. As tabelas de factos podem ser homogéneas ou heterogéneas, sendo possível para os mesmos dados, relativos a factos, coexistirem em tabelas diferentes.

Ferramenta de Reporting – software que, para além de outras funcionalidades, gera relatórios. Existem diversos produtos no mercado, incluindo a ferramenta usada neste projecto – o MicroStrategy.

Filtro – objecto que define condições lógicas que limitam o intervalo de dados numéricos apresentados num relatório. Limita os valores numéricos de um dado atributo a um intervalo previamente indicado.

Folha de Cálculo – documento que contém informações sumariadas de resultados obtidos destinados à análise de dados. Forma fácil de visualizar informações sobre um sistema e os seus componentes. As mais usadas são as folhas de cálculo Excel (Microsoft).

Hierarquia – forma de agrupar atributos de modo a espelhar as suas relações. O desenho de uma hierarquia consiste na organização de um grupo de atributos relacionados em áreas lógicas de negócio.

Metadados (ou Metadata) – definem-se como dados sobre dados. São dados que possuem a capacidade de descrever outros dados, ou seja, são a representação física de um objecto digital. No âmbito deste projecto, os metadados consistem num repositório cujos dados das tabelas de factos e dimensões do Data Warehouse, possibilitando o mapeamento dos conceitos de negócio. O repositório de metadados pode residir no mesmo servidor do Data Warehouse ou num servidor de Bases de Dados diferente.

Métrica – consiste num cálculo analítico obtido através de uma expressão construída a partir de objectos existentes, sendo estes: funções, atributos, factos ou outras métricas.

MicroStrategy – poderosa ferramenta de *Reporting* usada no âmbito deste projecto.

MicroStrategy Desktop – componente do MicroStrategy que proporciona a integração, a monitorização, a análise e o *Reporting* de suporte à decisão, num ambiente de desenvolvimento intuitivo. Em suma, permite aceder e gerir toda a componente de negócio de uma empresa.

MicroStrategy Enterprise Manager – componente administrativo do MicroStrategy onde se gere toda a informação através do *Reporting* e da consequente análise dos resultados produzidos. Este componente está inserido no ambiente MicroStrategy Desktop.

Microstrategy Intelligence Server – componente de gestão da prioridade dos pedidos de acesso ao Data Warehouse. Gestão de relatórios em *cache*, partilha de informação, protecção da *metadata* e dos objectos partilhados.

Microstrategy Narrowcast Server – servidor de entrega e distribuição de informação de negócio via *e-mail*, impressoras, serviços de ficheiros e dispositivos móveis. Através de uma arquitectura de distribuição robusta, cria um ambiente constituído por mecanismos de automatização, monitorização, análise de *performance* e *reporting*.

MicroStrategy Object – O termo *objecto*, no âmbito do MicroStrategy refere-se a atributos, factos, relatórios, métricas, filtros entre outros. A totalidade dos *objectos* é visível no MicroStrategy Desktop.

MicroStrategy Project – consiste no nível mais alto de intersecção entre um Data Warehouse, um repositório de metadados e os utilizadores/clientes e contém a totalidade dos *objectos* MicroStrategy. Tipicamente, sempre que se inicia uma sessão no MicroStrategy está a aceder-se a um projecto no qual estão definidos os *objectos* MicroStrategy correspondentes.

Modelo em Estrela (*Star schema*) – tipo de modelo dimensional, tipicamente usado num ambiente Data Warehouse. Consiste numa tabela de factos base relacionada com um qualquer número de tabelas de dimensões. Não existem relações entre dimensões.

Modelo em Floco de Neve – forma mais complexa do modelo em estrela. Estão presentes diversas tabelas de factos e um qualquer número de tabelas de dimensões agregadas a esta. As tabelas de dimensões possuem relações entre si. Adequa-se a maiores volumes de dados relacionados, mas introduz complexidade na exploração de dados.

Modelo em Constelação – forma mais complexa do modelo em estrela. Adequa-se a volumes, ainda maiores, de dados relacionados entre si. Recorre ao uso de um Bus – arquitectura de barramento de dimensões.

Online Analytical Processing System (OLAP) – consiste numa abordagem tecnológica para gerar respostas rápidas a consultas/pesquisas analíticas no âmbito do modelo dimensional. Deriva de uma adaptação do OnLine Transaction Processing (OLTP) a uma perspectiva de negócio e de *reporting*.

Online Transaction Processing (OLTP) – é uma forma de processamento de transacções conduzido através de redes de computadores, em tempo-real.

Operational Data Store (ODS) – grande repositório onde estão carregados todos os dados, devidamente convertidos, filtrados e validados, sem objectivo de retenção de histórico. Desenhada para alimentar e integrar dados de múltiplas fontes de forma a suportar e facilitar diversas operações. Os seus conteúdos são armazenados para, posteriormente, serem integrados no Data Warehouse.

Pivot Table (Tabela Pivot) – formato de representação de dados destinado à análise, contabilização, cruzamento e representação sumária dos mesmos. A sua criação é por *drag-and-drop* e permite exportar facilmente os dados depois da sua análise. Está presente no Excel (Microsoft) e noutras ferramentas de folhas de cálculo.

Reporting – estratégia de *Business Intelligence* que consiste na disponibilização da informação da empresa. Permite a visualização, exploração e formatação de dados, com fins de suporte à decisão e melhoria da estratégia de negócio. A informação é disponibilizada via *Web* ou no *Desktop* da ferramenta de *reporting*.

Relatórios (ou Reports) – consistem numa síntese e integração de dados cujos resultados servem um vasto conjunto de propósitos. Visualização de informação para ajudar o suporte à decisão e visão de negócio. Podem ser, também, de carácter administrativo, ou seja, resultam da análise dos dados como processo de gestão de informação interna sobre a qualidade dos dados. Consistem numa *query* SQL cujos resultados são apresentados numa forma gráfica, *user-friendly* através das ferramentas de *reporting*.

Sistema de Suporte à Decisão – no contexto de *Business Intelligence*, consiste num sistema que sustenta a tomada de decisões de negócio. Através dos resultados obtidos, é possível adoptar soluções apoiadas nas conclusões e resultados obtidos, com a consequente resolução de problemas.

Sistemas Operacionais – consistem nos principais fornecedores de informação de um Data Warehouse, Data Mart ou do ODS.

Software de Reporting – conjunto de ferramentas de *Reporting* disponíveis no mercado. O MicroStrategy é a ferramenta usada neste projecto mas existem muitas outras.

SQL – Linguagem de Questões Estruturadas. Consiste numa linguagem de pesquisa declarativa para bases de dados relacionais.

Tabela – na perspectiva de Sistemas de Informação, consiste no componente físico primário de um Data Warehouse / Base de Dados, constituído por colunas e dados dos mais variados tipos.

Tabela de Dimensões – tabela do modelo Data Warehouse físico, tipicamente o esquema em estrela ou floco-de-neve. Destinam-se a armazenar atributos respeitantes a dimensões.

.TXT – extensão de ficheiro em formato de texto;

View (ou Vista) – é uma entidade lógica. Consiste numa declaração SQL que é armazenada na base de dados no de sistema, o *tablespace*. Uma View pode possuir os todos ou alguns atributos de uma tabela existente numa Base de Dados;

XML – sigla que significa eXtended Markup Language, linguagem standard muito utilizada hoje em dia em diversos contextos, delimitada por tags definidas de forma não rígida, ao contrário da linguagem HTML, por exemplo.

Lista de Acrónimos

BI – **B**usiness **I**ntelligence
BD (ou **DB**) – **B**ase de **D**ados – **D**ata **B**ase
DASD – **D**irect **A**ccess **S**torage **D**evice
DBMS (ou **SGBD**) – **D**ata **B**ase **M**anagement **S**ystem – **S**istema de **G**estão de **B**ase de **D**ados
DDL – **D**ata **D**efinition **L**anguage
DML – **D**ata **M**anipulation **L**anguage
DCL – **D**ata **C**ontrol **L**anguage
DM – **D**ata **M**art
DSS (ou **SSD**) – **D**ecision **S**upport **S**ystems – **S**istemas de **S**uporte à **D**ecisão
DW – **D**ata **W**arehouse
ETL – **E**xtract, **T**ransform e **L**oad
HTML – **H**ypertext **M**arkup **L**anguage
IRS – **I**nformation **R**etrieval **S**ystems
MEI – **M**estrado em **E**ngenharia **I**nformática
MSTR – **M**icro**S**trategy
MIS (ou **SIG**) – **M**anagement **I**nformation **S**ystem – **S**istema de **I**nformação de **G**estão
ODS – **O**perational **D**ata **S**tore
OLAP – **O**nline **A**nalytical **P**rocessing **S**ystem
OLTP – **O**nline **T**ransaction **P**rocessing **S**ystem
PL/SQL – **P**rocedural **L**anguage/**S**tructured **Q**uery **L**anguage
SQL – **S**tructured **Q**uery **L**anguage
TI – **T**ecnologias de **I**nformação
XML – **e**Xtended **M**arkup **L**anguage

Capítulo 1

Introdução

Este capítulo introduz o tema do trabalho realizado. O Projecto, a motivação, as contribuições e resultados deste trabalho irão ser descritos numa perspectiva introdutória. O enquadramento institucional onde o projecto foi desenvolvido e a estrutura desta dissertação, estão também patentes neste capítulo.

1.1. Projecto em Engenharia Informática

A realização deste projecto insere-se no âmbito da disciplina de *Projecto em Engenharia Informática* do Mestrado em Engenharia Informática, segundo ciclo de Bolonha, da Faculdade de Ciências da Universidade de Lisboa. A área de especialização deste MEI é em Sistemas de Informação.

A realização deste Projecto em Engenharia Informática tem como objectivos gerais aprofundar e solidificar conhecimentos adquiridos a nível académico. Sendo efectuado numa empresa externa de acolhimento, proporciona a integração na vida profissional activa e a consequente adaptação a um ambiente típico de produção.

O desenvolvimento deste trabalho propicia o contacto com novas metodologias, ferramentas e responsabilidades.

A aprendizagem de conceitos técnico-científicos teóricos complementa a aquisição de conhecimentos práticos ao longo deste estágio.

Com o presente trabalho, pretende-se apurar a escrita de relatórios e de documentação técnica, o que fomenta o poder de síntese e a análise crítica de resultados atingidos.

Finalmente, tem-se como propósito aperfeiçoar a capacidade de apresentação pública dos resultados obtidos.

1.2. Motivação

Para uma organização, a informação consiste no seu bem mais precioso.

A nível empresarial, a qualidade da informação e a capacidade de responder, de modo flexível, às rápidas mudanças de ambiente consistem nos factores primordiais para o sucesso de uma área de negócio. As actividades quotidianas de tomada de decisão necessitam de grandes volumes de informação, orientada e organizada por assuntos.

No início da década de 90, os Sistemas de Data Warehouse (DW) foram emergindo como solução para suportar a estratégia de negócio – *Business Intelligence* – e o processo de suporte à tomada de decisão. Estes sistemas estão, hoje em dia, amplamente disseminados a nível empresarial. Têm contribuído para o bem-estar global das empresas uma vez que possuem a capacidade de armazenar e gerir os seus grandes volumes de informação de gestão. Possuem, ainda, uma grande capacidade de resposta a interrogações de gestão que necessitam de suporte temporal para a sua satisfação.

Dada a dinâmica dos mercados actuais, toda a área de negócio é confrontada diariamente com situações que, para além de requererem este grande volume de informação a nível da organização, exigem também a integração de informação contida em fontes externas à empresa.

A par destes constantes avanços tecnológicos está o crescimento exponencial da informação, bem como a urgência da tomada de decisões de negócio estratégicas, em curtos intervalos de tempo. Deste modo, a informação tem de ser constantemente gerida de forma a garantir a sua consistência, actualização e qualidade, traduzindo-se num factor de sucesso a nível empresarial.

Numa vertente arquitectural, os Sistemas DW estão intimamente ligados à tecnologia de Bases de Dados, lidando, com informação proveniente de diversas fontes externas. A informação é extraída, transformada e carregada, num processo de integração da mesma nos repositórios de informação da arquitectura física. Depois de integrada e armazenada esta informação é finalmente restituída aos utilizadores finais, de um modo completamente transparente. Esta restituição da informação é feita sob a forma de relatórios (*reports*) e consiste no ambiente típico de *Reporting* de uma área de negócios.

Este projecto tem como objectivo a análise e gestão de informação, com o intuito de melhorar o desempenho e a eficiência do ambiente de *Reporting*. Este trabalho pretende, igualmente, cumprir o requisito aplicacional de otimizar o espaço de armazenamento físico através da eliminação de objectos redundantes, obsoletos e sem utilização.

A metodologia adoptada neste projecto visa responder a um conjunto de problemas, inicialmente existentes, de modo a cumprir os objectivos propostos, num processo típico de Engenharia.

A combinação desta metodologia com ferramentas específicas, conceitos práticos e teóricos visa dar resposta aos problemas identificados em todos os níveis de agrupamento da informação, com enfoque na satisfação dos clientes/utilizadores finais.

Finalmente, todo este trabalho estará reflectido num Sistema de *Reporting* de gestão administrativa que permitirá rastrear soluções para prevenir a acumulação de informação ao longo do tempo. Deste modo, automatiza-se este processo de filtrar informação, evitando a sobrecarga do sistema de *Reporting*.

1.3. Enquadramento, Contribuições e Resultados

De modo a enquadrar as contribuições deste trabalho e o âmbito dos resultados obtidos faz-se uma breve panorâmica do ambiente subjacente ao projecto concretizado.

Esta dissertação gira em torno do ambiente DW de uma instituição bancária. A este ambiente associam-se conhecimentos ao nível de arquitecturas de referência, conceitos metodológicos e a exploração da potencialidade de novas ferramentas.

Nesta dissertação abordam-se os conceitos teóricos e práticos associados a um Sistema DW. No decurso deste projecto, a identificação de conceitos arquitecturais combinados com a exploração de informação de gestão permitiu um estudo minucioso que serviu de base para a definição do problema existente.

A procura de resolução para os problemas inicialmente identificados, resultou na definição de um conjunto de objectivos a cumprir, através da aplicação de uma metodologia bem definida. Este caso típico de Engenharia consistiu numa das principais metas a atingir neste projecto.

Este trabalho teve enfoque na análise e gestão de toda a informação presente num Repositório de Metadados com vista a reduzir, significativamente,

a quantidade de informação obsoleta existente. Estas práticas são conseguidas através da utilização de ferramentas especializadas, técnicas de exploração de informação e desenvolvimento de *Reporting* administrativo sobre a arquitectura DW existente. Toda esta metodologia de trabalho visa gerar um conjunto de resultados que derivam da aplicação de regras bem definidas. Todos os resultados obtidos foram alvo de avaliação com base nas medidas de precisão e abrangência dos mesmos.

Como contribuições a salientar da aplicação deste processo típico de Engenharia, resulta um ambiente com eficácia comprovada e com uma redução eficiente de espaço de armazenamento físico em memória.

Todo o trabalho anteriormente descrito serve a maior contribuição a evidenciar neste projecto. Esta contribuição consiste na preparação de todo o sistema de *Reporting* para um projecto futuro de grandes dimensões, que fará uso do ambiente organizado e de qualidade resultante deste trabalho.

Como principais contribuições salientam-se as seguintes:

- Produção de um Sistema automático de análise e gestão de informação presente num Repositório de Metadados, através da criação de um ambiente de *Reporting* administrativo de gestão;
- Identificação de um conjunto de objectos desactualizados ou sem utilização actual, através do estabelecimento de regras precisas;
- Possibilidade de reutilização do ambiente de *Reporting* de gestão administrativa concretizado. A concretização deste trabalho garante um modo automático de gerir a informação existente, identificando e solucionando problemas ao nível da desactualização e acumulação de informação ao longo do tempo;
- Avaliação dos resultados obtidos com base nas medidas de precisão e abrangência dos mesmos, o que permitiu comprovar a eficácia do Sistema de *Reporting* após a remoção de um conjunto de informação sem utilização actual, como resultado directo da aplicação da metodologia seguida neste projecto;
- Remoção efectiva de 2 884 objectos sem utilização;
- Redução de 27% do espaço de armazenamento físico inicialmente ocupado;
O que se traduz na concretização do requisito aplicacional de redução da carga ao nível do espaço físico de armazenamento em uso;
- Finalmente, com base em todo o trabalho desenvolvido e nos resultados obtidos, o presente projecto representa um ponto de partida para um projecto a realizar na instituição de acolhimento. O ambiente de *Reporting* conseguido através deste projecto caracteriza-se pela sua organização, pela existência de informação actualizada e pela redução de carga de armazenamento. Por tudo isto, servirá de base a um trabalho futuro de grande escala, que em tudo beneficiará destas características do ambiente de *Reporting*.

1.4. Enquadramento Institucional

Nesta secção é feito o enquadramento institucional no sentido de se caracterizar a natureza profissionalizante deste estágio.

Os aspectos relacionados com o processo de integração e as principais responsabilidades de um colaborador serão também alvo de reflexão.

1.4.1. Detalhes da Estrutura Organizacional

O desenvolvimento deste projecto na teve início a 1 de Setembro de 2006 com uma duração total de nove meses.

A empresa onde este projecto foi desenvolvido faz parte integrante de uma instituição bancária, sendo um grupo financeiro especialista em diversos sectores. Sustenta na rendibilidade e eficiência, perseguindo o objectivo de criação de valor.

A empresa onde este projecto se realizou é responsável pela gestão e coordenação de toda a área de tecnologias de informação e comunicação, possuindo uma organização interna por departamentos, como se representa esquematicamente. [Figura 1]

<Confidencial>

Figura 1 – Estrutura Organizacional da Área Informática do Grupo Bancário

Este projecto foi realizado num Departamento responsável pela Gestão Aplicacional, Exploração de Informação, Análise, Modelação e Implementação de Aplicações a utilizar pela totalidade do Grupo Bancário. A subdivisão do departamento em diversas áreas funcionais está representada em seguida. [Figura 2]

<Confidencial>

Figura 2 – O Departamento da Instituição Externa: Principais Competências

Todas as competências deste departamento estão directamente associadas a uma arquitectura de informação típica de um DW.

A arquitectura de informação existente é representada, de modo simplificado, de acordo com o esquema abaixo. [Figura 3]

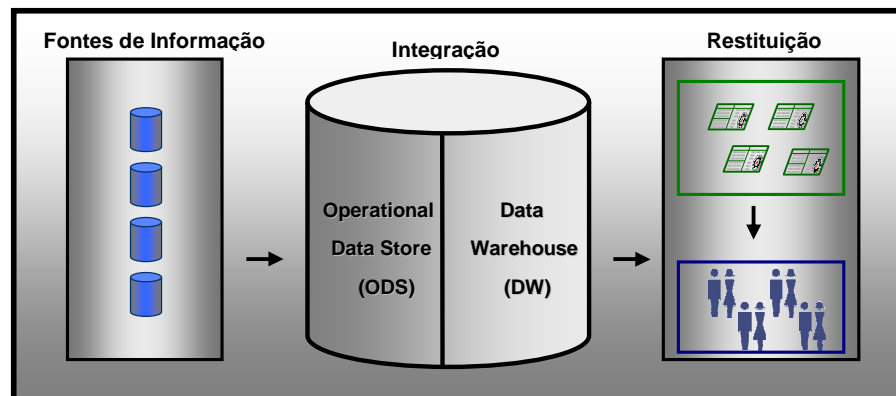


Figura 3 – Arquitectura da Informação do Departamento

Numa perspectiva resumida, esta arquitectura sintetiza todo o processo de integração da informação proveniente de fontes externas e a sua restituição aos utilizadores finais, consistindo num dos focos do ambiente de trabalho do Departamento.

1.4.2. A Integração na Instituição de Acolhimento Externa

<Confidencial>

1.4.3. Competências e Responsabilidades de um Colaborador

Em qualquer instituição, um colaborador tem como objectivo o cumprimento de um conjunto de responsabilidades bem definidas. Assim, o colaborador trabalha com base na competência, no rigor e no profissionalismo, garantindo, deste modo, um bom desempenho.

No âmbito deste projecto, salienta-se um conjunto de competências e metas a alcançar por um colaborador, listadas abaixo. [Tabela 1]

Responsabilidades e Competências de um Colaborador
▪ Sentido crítico na análise de dados e resultados obtidos;
▪ Cumprimento atempado do planeamento das tarefas;
▪ Postura inovadora na abordagem ao problema;
▪ Atitude positiva face ao projecto;
▪ Alinhamento entre os processos internos e de controlo do projecto;
▪ Postura proactiva e adaptável a diferentes ambientes;
▪ Usar a metodologia mais adequada para cumprir objectivos e garantir a resolução de problemas inicialmente identificados;
▪ Seguir criteriosamente cada uma das fases da metodologia;
▪ Asseverar a inexistência de erros no produto final entregue ao cliente.
▪ Assegurar a qualidade do produto final entregue ao cliente;
▪ Garantir a satisfação do cliente;

Tabela 1 – Responsabilidades e Competências de um Colaborador

No decurso deste projecto existiu uma procura contínua de aperfeiçoamento em que, para além do cumprimento dos objectivos, se aplicaram métodos de acordo com as necessidades do projecto. A busca de alternativas, o foco nas necessidades do cliente e a consolidação dos conhecimentos teóricos e práticos consistem nalgumas das principais práticas adoptadas no decurso deste projecto.

1.5. Estrutura da Dissertação

Após o capítulo introdutório, o presente documento está organizado nos capítulos que se mencionam em seguida.

O capítulo 2 apresenta o contexto técnico-científico do projecto, onde se caracterizam todos os conceitos de DW, *Business Intelligence* e *Reporting* subjacentes ao projecto.

O capítulo 3 contempla os objectivos e a metodologia que garante o cumprimento das metas propostas. As tarefas inicialmente delineadas e o planeamento efectivamente concretizado são também descritos e comparados neste terceiro capítulo.

No capítulo 4 faz-se uma descrição pormenorizada de todo o trabalho realizado e das contribuições daí decorrentes. Faz-se uma apresentação das ferramentas e tecnologias utilizadas. Apresentam-se os resultados obtidos e a sua avaliação com base na precisão e abrangência. Efectuam-se análises críticas às ferramentas, tecnologias, trabalho e resultados obtidos.

No capítulo 5 estão patentes a conclusão e o trabalho futuro que tem como base o trabalho desenvolvido neste projecto.

Capítulo 2

Panorâmica

Neste capítulo são apresentados os conceitos necessários à compreensão de um ambiente *Data Warehouse* (DW) típico. É feita uma panorâmica geral das tecnologias DW enquadrando-as no âmbito deste projecto. Nesta secção, aborda-se a forma como estas tecnologias representam uma mais-valia no contexto empresarial. É feita uma caracterização do conceito DW, descrevendo-se cada uma das visões distintas que este pode assumir. Os conceitos primordiais para a percepção de um sistema de *Reporting* são igualmente alvo de atenção. Este capítulo contempla ainda a evolução histórica do conceito de *Reporting*.

2.1. Panorâmica Geral

Desde o início da década de 90 que as tecnologias DW estão amplamente disseminadas ao nível das organizações. Estes sistemas emergiram de forma a garantir a melhoria do processo de tomada de decisão e o apoio às estratégias de mercado, contribuindo para o bem-estar geral das empresas. Esta contribuição advém da capacidade destes sistemas armazenarem e gerirem grandes volumes de informação de gestão e da sua capacidade de resposta rápida, em tempo-real.

Esta dissertação aborda o problema do crescimento exponencial da informação armazenada que é continuamente alterada, num ambiente DW. A grande capacidade de armazenamento e o processo de contínuo refrescamento da informação, embora sejam mais-valias, a longo prazo poderão resultar na acumulação de um conjunto de informação duplicada, desactualizada ou mesmo obsoleta. Todas estas circunstâncias poderão colocar em risco o desempenho geral do Sistema de Informação.

O conceito de *Reporting*, tema fulcral neste projecto, traduz-se na restituição de informação aos clientes através da construção e entrega de relatórios com toda a informação por estes requisitada. Novamente, surge o problema da acumulação de informação que poderá comprometer o sistema, na entrega de informação de negócio, sob a forma de *Reporting*.

Para atenuar este problema, carece-se de um processo complexo de análise e gestão da informação ao nível do repositório de dados onde residem as tabelas presentes na arquitectura DW, reflectindo os conceitos de negócio.

Em suma, esta dissertação explora o problema da acumulação de informação ao longo do tempo. A análise e gestão do conhecimento exigem um profundo estudo dos conceitos teóricos e arquitecturais subjacentes a este ambiente DW que se irão descrever de seguida, em pormenor.

2.2. O Conceito Data Warehouse

O conceito Data Warehouse (DW) tem sido alvo de definições e interpretações distintas ao longo do tempo.

Destacam-se as visões dos dois autores mais conceituados nesta área:

- **Bill Inmon's paradigm:** *"Data warehouse is one part of the overall Business Intelligence system. An enterprise has one data warehouse, and data marts source their information from the data warehouse. (...) used to support the business' decisions."* [18]
- **Ralph Kimball's paradigm:** *"Data warehouse is the conglomerate of all data marts within the enterprise. Information is always stored in the dimensional model."* [6]

Os autores destacam as visões que o conceito DW pode tomar. Não existe nenhuma perspectiva correcta ou errada, apenas representam um modo peculiar de olhar para um DW.

De um modo geral, a perspectiva de Kimball aproxima-se mais das necessidades de informação em termos empresariais, focando, principalmente aspectos ao nível da arquitectura de referência – o modelo dimensional. Inmon é considerado o pai do DW, este autor salienta a necessidade de suporte às decisões de negócio. [16] No entanto, ambos os autores concordam que o sucesso de um sistema DW depende de um levantamento correcto dos requisitos de negócio.

O DW assume um carácter filosófico, podendo ser interpretado de acordo com diferentes perspectivas. O conceito DW desdobra-se, assim, em duas perspectivas principais. Uma diz respeito a detalhes arquitecturais e a outra direcciona-se para uma perspectiva metodológica no âmbito das áreas de *Business Intelligence* e suporte à decisão/análise de negócio.

2.2.1. Perspectiva Arquitectural

A perspectiva arquitectural consiste na definição de um modelo arquitectural de referência que suporta o processo de extracção, transferência e carregamento da informação (ETL).

Nesta perspectiva, um DW consiste numa colecção de dados integrados, não voláteis e orientados a um assunto. Existe o conceito de histórico, ou seja, os dados são registados ao longo de um período temporal bem definido, de acordo com as necessidades do período de armazenamento da informação.

A definição de um modelo arquitectural que suporte todos os componentes necessários ao processo de ETL, assenta na existência de diversos componentes que quando relacionados, resultam na restituição de toda a informação, depois de integrada, aos utilizadores finais. Esta restituição de informação é efectuada de modo completamente transparente para os utilizadores/clientes, ocultando todo o processo de transformação e integração de informação nas respectivas estruturas de dados.

Esta arquitectura de referência de um DW pressupõe um fluxo de informação que contempla o DW e um conjunto de componentes correlacionados.

A base deste sistema são as fontes externas de informação que "alimentam" o Operational Data Store (ODS) através de processos ETL que tratam os dados a "injectar" no DW e/ou nos Data Marts. Finalmente, a informação é armazenada para que depois possa ser organizada sob a forma de *Reporting* e restituída aos utilizadores/clientes finais.

A definição do fluxo de informação subjacente à arquitectura DW será, em seguida, complementada com uma panorâmica do conjunto de componentes de um DW, de modo a enquadrar o processo final do fluxo de informação, a restituição de informação, sob a forma de *Reporting* aos clientes/utilizadores finais.

- **Fontes Externas ou Sistemas Operacionais**

Consistem nas fontes iniciais que alimentam/fornecem a informação ao ODS e, conseqüentemente, ao DW. São fornecidas em “bruto” ou por aplicações, chegando em diversos formatos, por exemplo ficheiros “.txt”, Excel ou outros.

- **Extraction, Transformation and Load (ETL)**

Processo de Extração, Transformação e Carregamento de dados.

Este processo é iniciado com a extração de dados provenientes de fontes externas, maioritariamente ficheiros provenientes de outras aplicações. Seguem-se processos de transferência e tratamento dos dados. Finalmente, os dados podem ser carregados de modo a que a informação esteja preparada para ser restituída aos utilizadores finais.

Ao nível da arquitectura, este processo pode ser utilizado para popular não só o DW como também o *Operational Data Store* e os Data Marts.

- **Operational Data Store**

O *Operational Data Store* (ODS) consiste no repositório de informação de natureza operacional de uma organização. É uma estrutura de armazenamento volátil e altamente detalhada, usada para consolidar todas as informações sobre um determinado assunto, muitas vezes dispersas nos vários sistemas da empresa. Pode possuir somente dados “em bruto” vindos das fontes externas ou armazenar dados que já sofreram algum tipo de tratamento (limpeza, consolidação de chaves, etc.).

Se possuir informação histórica, esta é muito reduzida, uma vez que os dados históricos são guardados no DW. Deste modo, o ODS serve de intermediário para os dados que vão para o DW.

O ODS serve de ligação dos Sistemas Operacionais ao DW e aos Data Marts.

As vantagens desta abordagem arquitectural são a melhoria da qualidade das decisões tácticas proporcionada pelo ODS e a optimização do processo de extração e carregamento dos dados no DW.

A implementação desta abordagem também pode ser longa e dispendiosa, requerendo um planeamento bem detalhado.

O ODS opera directamente conectado aos dados operacionais ou fontes de informação, objectivando o suporte a decisões de natureza operacional, com características que permitem a obtenção de tempos de resposta bastante rápidos, algo que um DW clássico não consegue prover.

▪ **Data Mart**

Consiste num componente de um DW orientado a um objectivo muito específico. Normalmente surgem para tratar, de forma especializada, um conjunto de informação de um DW.

Em termos arquitecturais, podem ser representados de modo independente ou federado:

- **Data Marts Independentes:** armazenam toda a informação que utilizam, funcionando como um DW de menor tamanho;
- **Data Marts Federados:** partilham um conjunto de informação com o DW ou com outros Data Marts, podendo estar integrados no próprio DW.

▪ **Exploração de Informação**

A exploração de informação permite a análise das estruturas de dados. No âmbito deste projecto, a exploração de informação foi conseguida através do uso de ferramentas e programação em PL/SQL.

▪ **Restituição de Informação**

A restituição de informação assenta na produção de um conjunto de relatórios entregues aos clientes da área de negócio. Um sistema de *Reporting* consiste na entrega de conhecimento a clientes/utilizadores finais, de acordo com as necessidades de cada uma das áreas existentes. Esta restituição da informação ao cliente ocorre de um modo completamente transparente, ocultando todo o processo de extracção, carregamento, transformação e integração da informação.

▪ **Metadata**

O conceito de Metadata (Metadados) é um conceito primordial, neste projecto. Consiste num repositório de dados onde as tabelas de um DW estão mapeadas, reflectindo os conceitos de negócio. O repositório de metadados pode residir no mesmo servidor do DW ou num servidor de Bases de Dados diferente.

Os metadados possibilitam a definição de processos de gestão e manipulação dos dados sobre o DW. Os diferentes tipos de metadados englobam as especificações das fontes de informação, a aquisição de dados, o dimensionamento das tabelas de origem e de destino no processo de migração dos dados, as definições dos próprios dados, entre outros.

Um DW típico e os seus componentes podem ser vistos como se mostra na seguinte arquitectura de referência. [Figura 4]

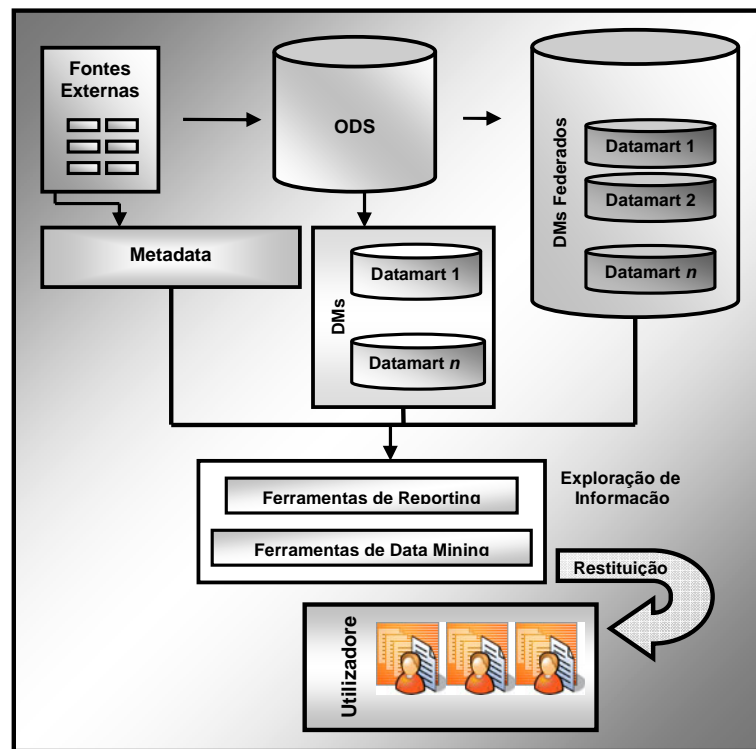


Figura 4 – Arquitectura Data Warehouse de referência

Esta arquitectura de referência é a subjacente a este projecto. Tanto o DW como os componentes típicos anteriormente descritos se adaptam à realidade existente no Departamento onde este trabalho foi desenvolvido.

2.1.1.1. Modelação Dimensional

Neste projecto, a análise e a gestão da informação assumem um papel da maior importância, cuja eficiência depende do conhecimento da estrutura dos mesmos.

O Modelo Dimensional, usado no ambiente DW subjacente a este projecto, caracteriza-se pela organização, num arranjo peculiar, que traduz as relações ao nível de tabelas distintas do modelo. Como componentes salientam-se factos, dimensões e respectivas ligações.

Em seguida define-se cada um dos conceitos que integram o modelo dimensional. [Tabela 2]

Componentes do Modelo Dimensional	Definição
Factos	Consistem em registos de ocorrências, tendem a ser mensuráveis, sendo muitas vezes referidos como medidas. Possibilitam a análise a diferentes níveis de granularidade e tipicamente são numéricos e aditivos. É imprescindível que possam ser agregados (funções de agregação).
Dimensões	Conjunto de atributos que descrevem uma mesma propriedade da ocorrência. Tipicamente, os atributos de uma dimensão são descritivos e não mensuráveis (valores textuais com um número limitado de possibilidades). São organizados segundo hierarquias bem definidas. Possibilita a análise dos factos a diferentes níveis de granularidade. Fornecem as restrições mais interessantes para análise. Se os factos registam as ocorrências de um negócio, as dimensões descrevem as condições do negócio.
Tabela de Factos (Factual)	Contêm um ou mais atributos que ocorrem numa dada combinação de registos das várias dimensões. É esta combinação de registos que identifica univocamente cada facto, através do conceito de chave. Traduz sempre uma relação de muitos para muitos, possuindo uma chave primária composta por múltiplas chaves estrangeiras (uma para cada dimensão que caracteriza a ocorrência do facto). [5]
Tabela de Dimensões	Contêm os diferentes atributos respeitantes à dimensão e um identificador unívoco para cada registo da tabela, permitindo, assim a ligação à Factual. [5]

Tabela 2 – Componentes do Modelo Dimensional

O Modelo Multidimensional representa o registo de dados em tabelas, relacionando factos e dimensões. Esta forma de modelar a informação pode apresentar diferentes arranjos, denominados esquemas. De acordo com a complexidade das relações existentes entre tabelas de factos e tabela de Dimensões, os esquemas apresentam diferentes configurações, sendo estas:

- Esquema em Estrela;
- Esquema em Floco de Neve;
- Esquema em Constelação.

O esquema em estrela é o que apresenta o desenho mais simples. Existe uma tabela de factos ligada a um conjunto de tabelas de dimensão, não existindo quaisquer relações entre as tabelas de dimensões. Apresenta-se esquematicamente um modelo em estrela de referência. [Figura 5]

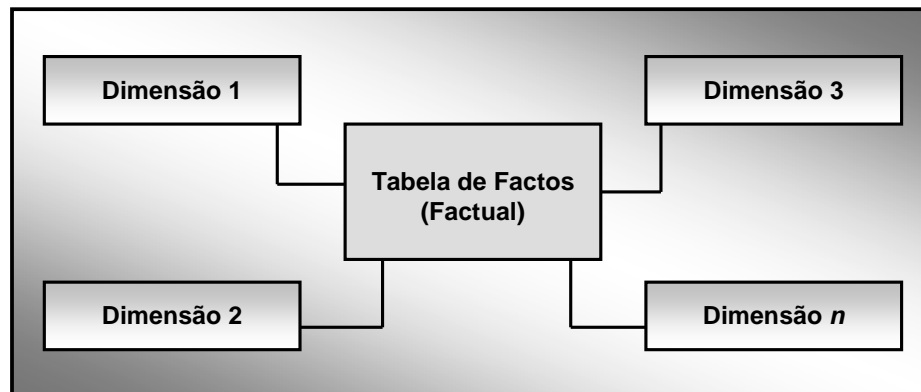


Figura 5 – Esquema em Estrela

O esquema em floco de neve é um refinamento do esquema em estrela. Cada dimensão pode ser representada por mais do que uma tabela de dimensão, existindo relações entre tabelas de dimensões distintas. Este modelo não é muito utilizado dado que introduz complexidade na exploração dos dados. [Figura 6]

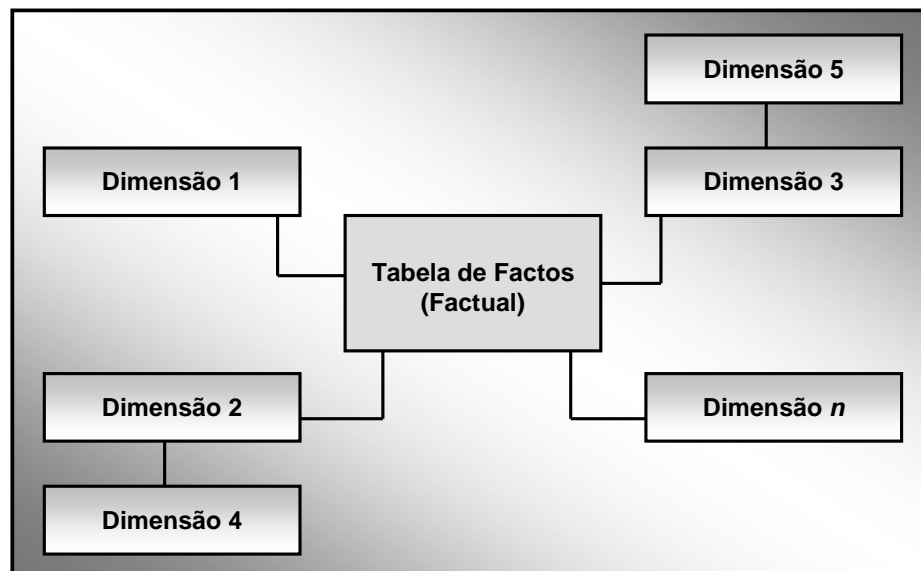


Figura 6 – Esquema em Floco de Neve

A configuração mais complexa consiste no esquema em constelação de factos que se representa por um conjunto de esquemas em estrela ou em floco de neve ligados entre si, por um conjunto de relações. Tal como o esquema em floco de neve, é pouco utilizado, dado que a complexidade das relações torna a exploração de dados e a percepção do modelo muito complexas.

Embora se tivessem mostrado arquitecturas e modelos dimensionais padrão, existem muitos modos alternativos de desenho e modelação dimensional. Estes variam consoante o problema, as necessidades de modelação existentes, o tipo de informação a tratar e o produto final a desenvolver.

Mais recentemente, a modelação dimensional tem caminhado no sentido de se tornar multidimensional. A estruturação da informação é efectuada em cubos que agregam a informação por níveis hierárquicos. No âmbito deste projecto utilizou-se somente modelação dimensional.

2.2.2. Perspectiva Metodológica de Negócio

Na perspectiva metodológica, o conceito DW assume-se como uma estratégia que proporciona um ambiente favorável à tomada de decisões, numa área de negócios. A metodologia DW, na perspectiva de negócio, está directamente associada aos conceitos de *Business Intelligence* e Sistemas de Suporte à Decisão.

O conceito de *Business Intelligence* diz respeito à gestão de negócio, usando um conjunto de mecanismos que permitem tomar decisões de negócio em curtos espaços de tempo, visando uma estratégia que permite à empresa alcançar uma vantagem competitiva, face à concorrência. [24] A extracção do conhecimento, combinada com Sistemas de Suporte à Decisão e *Reporting* proporciona um ambiente tático que facilita a análise e gestão de todas as informações de mercado de uma organização.

Esta combinação de conceitos tem vindo a ser adoptada por imensas empresas dado o sucesso comprovado dos resultados produzidos, ao longo do tempo.

Para uma organização, o tipo de decisões a suportar centra-se, sobretudo:

- Em decisões estratégicas de mercado a longo prazo;
- Na análise da situação de mercado actual e planeamento da tática de melhoria de resultados futuros;
- Na decisão da forma de penetração no mercado existente;
- Na urgência de tomada de decisão, em curtos espaços de tempo;
- Na análise dos factores que afectam positiva ou negativamente o negócio;
- Em técnicas de Inovação.

Toda esta metodologia produz resultados sob a forma de relatórios financeiros. Estes documentos são restituídos aos clientes de modo completamente transparente, ocultando todos os detalhes de extracção e integração da informação. Em termos de trabalho relacionado, esta devolução de informação aproxima-se de um sistema típico de restituição de informação - *Information Retrieval Systems* (IRS). Estes sistemas são alvo da aplicação de metodologias de avaliação, dos resultados produzidos, em termos de precisão e abrangência dos mesmos. Deste modo é possível obter conclusões acerca da eficiência do sistema de *Reporting*.

2.2. O Conceito de *Reporting*

Num ambiente típico de *Reporting*, produzem-se relatórios com as mais variadas informações de negócio.

Qualquer ambiente de *Reporting* pode ser visto como um sistema de suporte à decisão. O objectivo comum a todos os utilizadores é receber relatórios (*reports*) com resultados que auxiliem e apoiem a tomada de decisão, resultando em sucessos na área de negócio.

2.3. *Reporting* – Evolução histórica

Desde a introdução do computador como ferramenta de trabalho, que são gerados ficheiros fonte e relatórios em papel.

A par dos avanços tecnológicos, estava a produção de ficheiros operacionais que serviam de base para a geração de relatórios em papel. O número de ficheiros tendeu a crescer exponencialmente, resultando em problemas ao nível da distribuição e manutenção dos dados. Outros problemas surgiam ao nível do armazenamento e do sincronismo dos dados. O processo de inserção de dados era efectuado manualmente e seguindo especificações pré-definidas, de acordo com a periodicidade dos relatórios. Findo este processo, os relatórios eram enviados para os seus requerentes, num processo estandardizado e geralmente muito moroso.

O aparecimento das Bases de Dados, na altura centralizadas como uma fonte única de dados para todo o processamento, consistiu num grande avanço tecnológico.

Mais tarde, a noção do conceito de transacção e o seu processamento consistiu noutro marco histórico no processo de *Reporting*.

Na década de 80, os avanços tecnológicos e o surgimento do computador pessoal fizeram com que novas práticas de trabalho e conceitos emergissem e se fossem consolidando.

Em seguida apresenta-se, esquematicamente, a evolução histórica do processo de *Reporting*. [Tabela 3] [18]

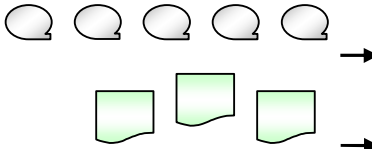
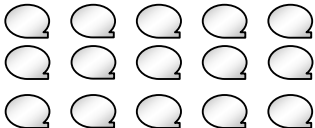
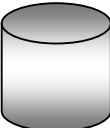
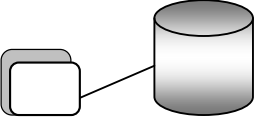
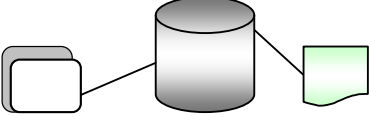
Data	Esquematização	Fontes e Conceitos	Vantagens e Desvantagens
1960		Ficheiros Fonte (Master Files)	Vantagem: <ul style="list-style-type: none"> Disponibilização Desvantagem: <ul style="list-style-type: none"> Distribuição e Tratamento
		Relatórios em Papel	Desvantagem: <ul style="list-style-type: none"> Sem suporte computacional
1965		Ficheiros Fonte (Master Files) Em grande número	Desvantagens: <ul style="list-style-type: none"> Complexidade ao nível de Manutenção e Desenvolvimento; Sincronização de dados; Hardware.
1970		Avanços no Modo de Armazenamento Sistemas de Gestão de BD	Vantagem: <ul style="list-style-type: none"> Base de Dados Desvantagem: <ul style="list-style-type: none"> Visão na época era uma fonte única de dados para todo o processamento. Centralização.
1975		Transacções	Vantagens: <ul style="list-style-type: none"> Online Boa Performance Processamento de Transacções
1980 a 1990		Gestão de SI Sistemas de Suporte à Decisão	Vantagens: <ul style="list-style-type: none"> Avanços Tecnológicos Computador Pessoal (PC)

Tabela 3 – Perspectiva Histórica

Nos últimos anos, a par da contínua evolução da tecnologia, surgiram ferramentas no âmbito de *Business Intelligence*, *Reporting* e suporte à tomada de decisão que auxiliam este processo de restituição de informação de negócio, sob a forma de relatórios. A par destas evoluções, o conceito DW foi-se consolidado, sendo hoje utilizado em larga escala a nível empresarial.

Sumário

Neste capítulo foram apresentados os principais conceitos associados a um DW, de modo a fornecer uma panorâmica geral deste ambiente e da sua inclusão no âmbito deste projecto. Destacaram-se as diferentes visões teóricas de um sistema DW. Salientou-se a perspectiva arquitectural contemplando os seus diversos componentes e a visão metodológica de negócio. Focou-se também o conceito de *Reporting* e a sua evolução tecnológica, ao longo do tempo.

Capítulo 3

Objectivos, Metodologia e Planeamento

Neste projecto os objectivos identificados derivam de problemas existentes. Para tal, existe a necessidade de encontrar soluções e concretiza-las de acordo com uma metodologia adequada. Este é um processo típico de Engenharia.

No presente capítulo é feito o enquadramento dos objectivos a satisfazer, face aos problemas e necessidades actuais. Faz-se uma descrição detalhada da metodologia adoptada, durante o ciclo de vida do projecto, visando o cumprimento dos objectivos. Finalmente, apresenta-se o planeamento inicial de tarefas e a sua confrontação com o plano de tarefas efectivamente concretizado.

3.1. Objectivos

O objectivo central deste projecto consiste na identificação de problemas e na procura de soluções, num caso típico de Engenharia. Estes problemas reflectem necessidades em diversos níveis, como se esquematiza na figura seguinte. [Figura 7]

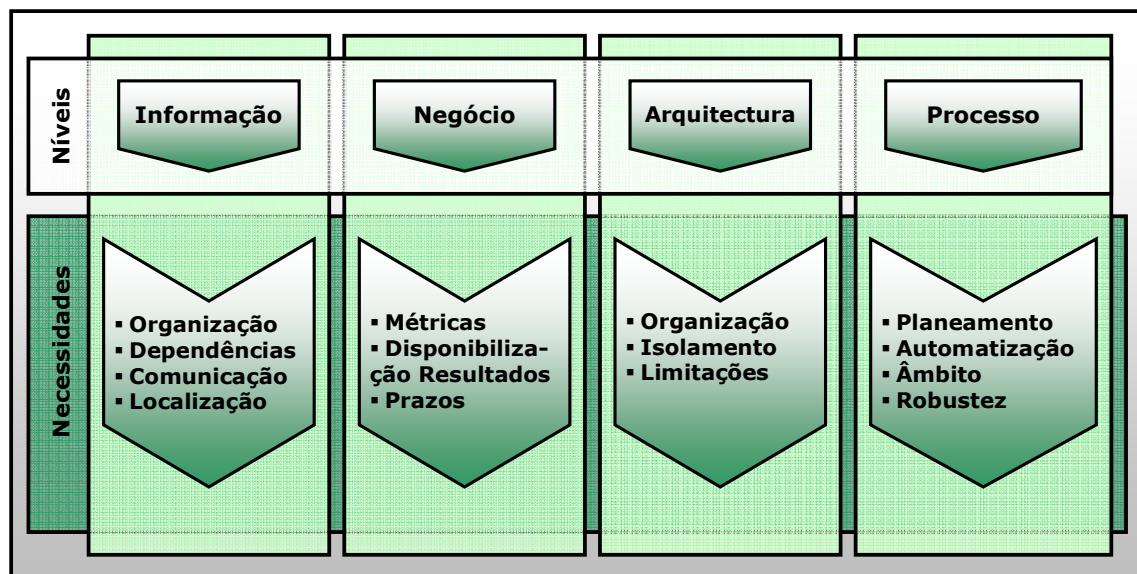


Figura 7 – Necessidades existentes em diferentes níveis

O bem mais precioso de uma empresa é a sua informação. Este nível abarca um conjunto de dados a analisar e gerir com vista a atingir resultados precisos e consistentes que comprovam o cumprimento de objectivos e a satisfação dos utilizadores finais.

Na perspectiva de negócio, a especificação de objectivos garante a obtenção de resultados que suportem decisões num prazo preestabelecido. Este nível centra-se entre o sistema e os utilizadores da área de negócio. Define-se como um modo de transformar necessidades e dificuldades numa solução óptima. Os problemas deste nível estão relacionados com a gestão, manutenção e robustez do sistema.

A identificação de problemas a nível arquitectural ocorre quando é necessário adaptar o sistema base existente às necessidades de comunicação e identificar possíveis limitações. Neste projecto, este nível é da maior importância dado que toda a gestão, análise e exploração de informação dependem directamente do contexto arquitectural ou estrutural da mesma. Torna-se, assim, necessário um estudo meticoloso dos componentes arquitecturais do sistema.

Ao nível dos processos existentes carece-se de uma breve análise dos processos automáticos, sistemáticos e das dependências existentes. Os problemas associados a este nível prendem-se com questões de robustez, administração e gestão.

Para satisfazer os objectivos, propõe-se uma solução que se descreve na próxima subsecção.

3.1.1. Solução Proposta em Diversos Níveis

A solução proposta consiste numa abordagem segmentada a diferentes níveis, que embora distintos, estão intrinsecamente ligados, como se apresenta na seguinte figura. [Figura 8]

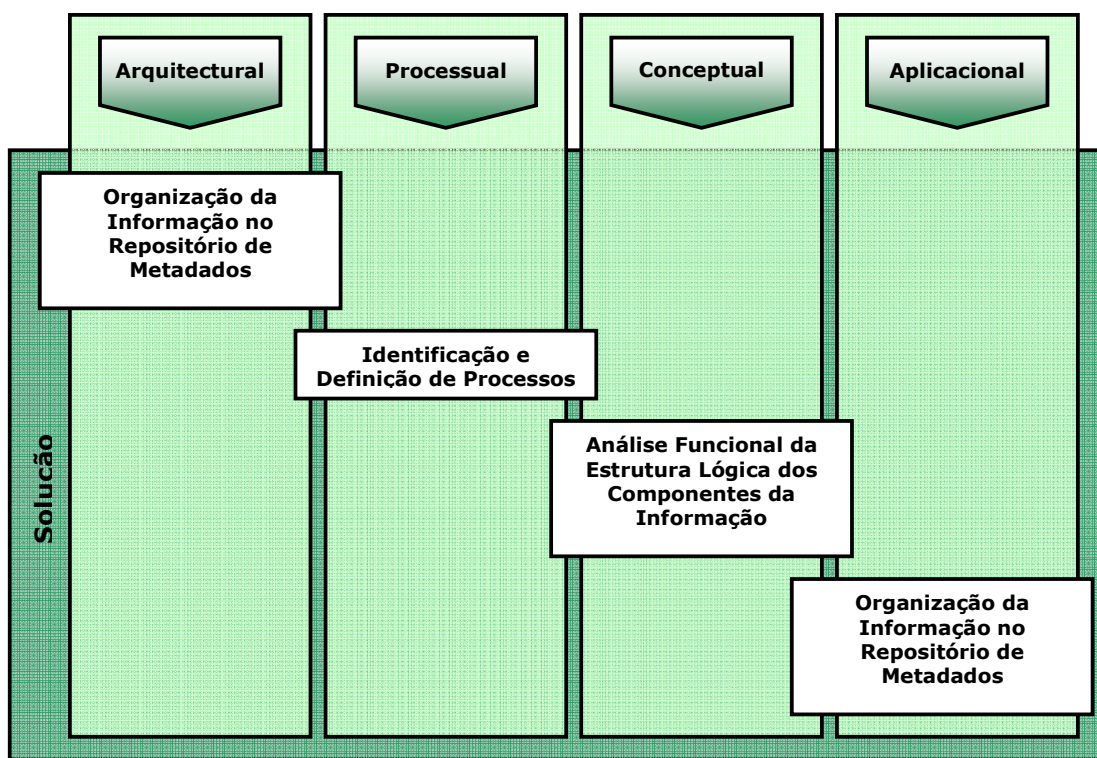


Figura 8 – Níveis da Solução proposta

Estes estágios cobrem a solução proposta nos diversos níveis que ocupam. Esta subdivisão organiza os componentes arquitecturais, os processos, os conceitos e a aplicação que fazem a ligação entre os dados num Sistema DW, reflectindo-se ao nível do *Reporting*.

Em seguida apresenta-se cada, com detalhe, cada um dos níveis da solução proposta.

3.1.1.1. Architectural

A arquitectura base deste sistema é bem organizada e encontra-se distribuída por diversas máquinas, ao nível físico.

Dada a especificidade deste projecto, não se considera relevante descrever em pormenor a infra-estrutura física subjacente ao sistema, motivo pelo qual se coloca somente em anexo. [Anexo 2]

As componentes DW e *Metadata* são aquelas que, neste projecto, foram totalmente cobertas, representadas no esquema abaixo. [Figura 9]

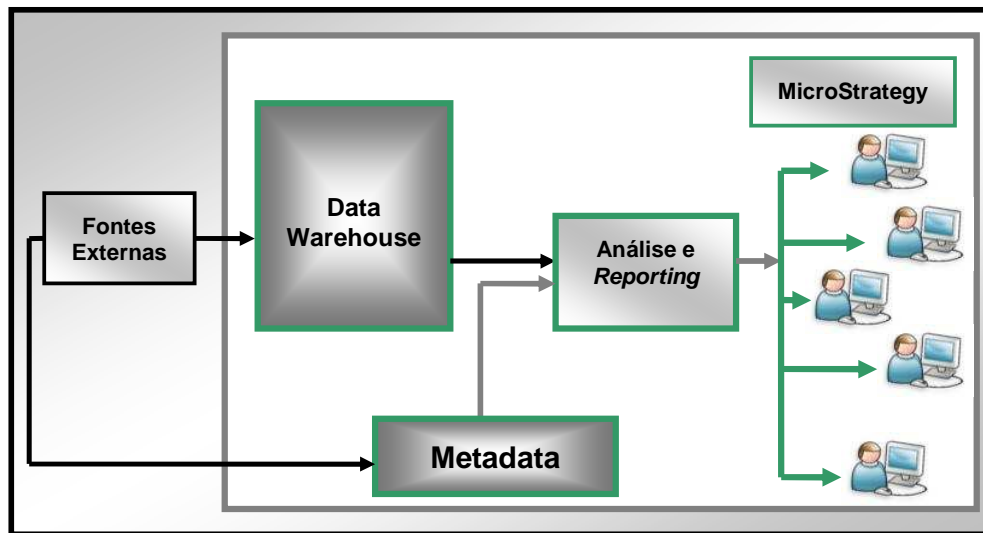


Figura 9 – Arquitectura Associada ao Projecto

Alimentados pelas fontes externas, os repositórios DW e Metadata possuem a informação que irá ser restituída aos utilizadores finais, sob a forma de relatórios (*reports*). A figura representa a arquitectura associada ao projecto, focando os aspectos da arquitectura de referência de um DW com maior relevo no âmbito deste trabalho. Está também patente a restituição da informação aos utilizadores, via ferramenta, sob a forma de *Reporting*.

Numa vertente lógica de arquitectura de informação, podemos considerar como componentes de maior relevância os que se descrevem abaixo. [Tabela 4]

Sistema de Reporting	Sistema de construção e distribuição de relatórios com informação de negócio, de modo a suportar a tomada de decisão dos utilizadores finais/clientes.
DW	Repositório de Dados com toda a informação de negócio. Pode assumir diferentes caracterizações, orientadas a uma visão arquitectural ou metodológica.
Ferramentas de Reporting	<ul style="list-style-type: none"> ▪ Ferramenta de Reporting; ▪ Construção de Relatórios; ▪ Distribuição de relatórios aos utilizadores finais via Web e através do ambiente da própria ferramenta – Desktop. <p>Possui um componente administrativo que possibilita a:</p> <ul style="list-style-type: none"> • Criação de Relatórios administrativos de gestão; • Criação de Relatórios com informação estatística.
Metadata	<ul style="list-style-type: none"> ▪ Repositório de metadados com toda a informação estatística e relativa a cada um dos objectos existentes: relatórios, tabelas, utilizadores, sessões entre muitos outros. ▪ Possui uma camada abstracta que correlaciona os objectos MicroStrategy.

Tabela 4 – Arquitectura Existente e Funcionalidades

Neste componente arquitectural, pretende-se explorar a informação existente num processo analítico. Como fim último visa-se a diminuição, em termos de armazenamento, o espaço físico ocupado por objectos obsoletos ou sem utilização recente. Para tal, é necessária uma análise de impactos para que se consigam detectar dependências existentes e constrangimentos no processo de remoção.

3.1.1.2. Processual

Todos os processos existentes neste ambiente DW deste projecto foram planeados para aumentar o desempenho, a fluidez e a coerência do processo global de *Reporting*.

Os processos já construídos foram explorados e analisados no sentido de identificar dependências entre objectos.

No âmbito deste projecto os processos desenvolvidos não foram ao nível do tratamento de dados mas sim ao nível da análise e gestão da informação de carácter aplicacional. O uso de componentes de administração da ferramenta de *Reporting* faculta, neste projecto, o processo de especificação de regras que possibilitam a detecção de objectos obsoletos ou sem utilização recente.

Foram também criados processos de filtragem de informação de acordo com regras de periodicidade, de modo a caracterizar a temporalidade de refrescamento e/ou actualização dos objectos existentes no repositório de Metadados.

3.1.1.3. Conceptual

Ao nível dos conceitos, neste projecto foi necessária a exploração da estrutura de dados existente, de modo a que posteriormente pudessem ser detectadas particulares que caracterizassem os objectos como sendo obsoletos, duplicados ou sem utilização recente. O conhecimento decorrente desta análise e exploração de informação, torna possível a criação de regras e métricas que se irão reflectir ao nível dos processos e da aplicação de *Reporting*.

A estruturação lógica da informação e a identificação de fluxos serve de base a todo o trabalho desenvolvido neste projecto.

A metodologia de trabalho concretizada teve como base todo este componente conceptual, cruzando todos os outros componentes: o arquitectural, o processual e o aplicacional.

O sistema concretizado assentou nesta vertente conceptual dado que a informação extraída tem por base camadas lógicas de informação em diferentes níveis conceptuais.

Numa fase final, os conceitos principais foram agrupados de modo a atingir a solução pretendida.

Em suma, ao nível conceptual foi necessária a implementação de uma metodologia de trabalho que possibilitou a estruturação lógica da informação e do fluxo de dados existentes, com vista à reorganização e eliminação de objectos e dependências obsoletas ou sem utilização. As diversas camadas lógicas e semânticas foram exploradas com grande detalhe de modo a organizar a informação proveniente do repositório de Metadados e do DW, criando níveis de abstracção que permitem estruturar a definição do restante trabalho a implementar.

3.1.1.4. Aplicacional

A aplicação desenvolvida consiste num acréscimo de competências ao sistema de *Reporting* já existente. Para além disso, este projecto serve de base a uma outra aplicação, de maiores dimensões, a concretizar de futuro.

Neste nível foram criados relatórios de gestão, com carácter estatístico que devolvem a informação pretendida, depois de terem sido estabelecidas métricas e regras sobre os dados. Numa metodologia modular, todo o processo de análise e gestão é feito em conjunto e revisto pelos utilizadores, sendo alvo de uma segunda análise de pertinência.

A informação aplicacional, representada como ligação dos utilizadores às estruturas de dados do componente arquitectural, é disponibilizada tendo em vista um ambiente organizado, mais eficiente e caracterizado pela qualidade. A remoção de informação obsoleta visa alcançar o ambiente anteriormente descrito, bem como a diminuição da carga do sistema de *Reporting*. A reutilização deste sistema de *Reporting* administrativo de gestão é também um propósito a garantir, fazendo com que o processo de detecção de informação obsoleta seja automatizado, de futuro.

Esquemáticamente, a componente aplicacional pode ser representada de acordo com a figura abaixo. [Figura 10]

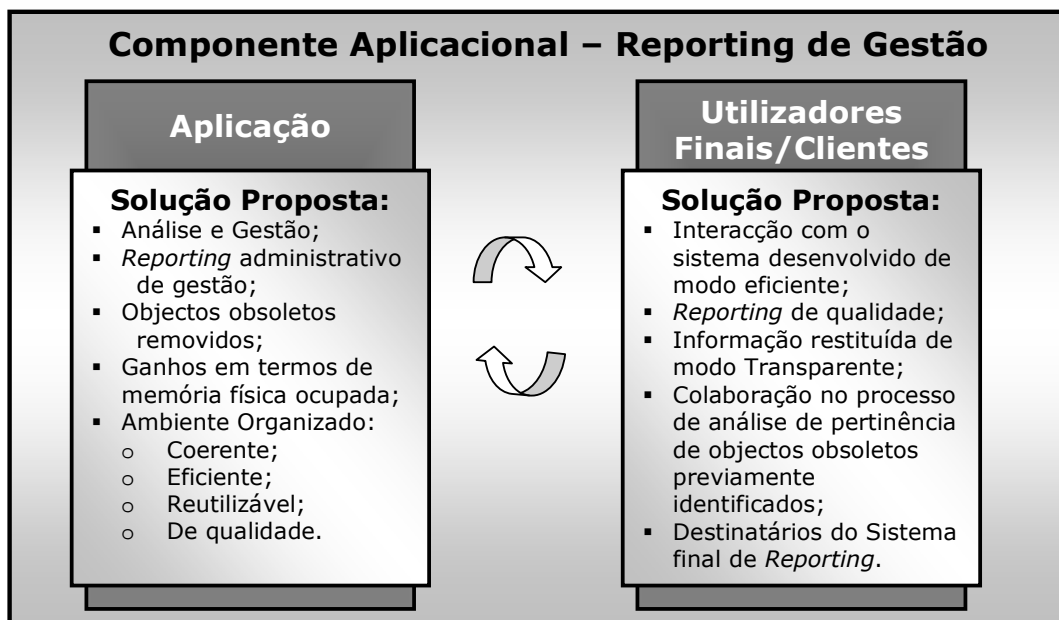


Figura 10 – Relação Aplicação – Utilizador Final/Cliente

Em suma, a solução proposta para o problema existente abarca diferentes níveis que, quando interligados, permitem a definição da metodologia de trabalho a seguir para atingir os objectivos, inicialmente propostos para este projecto.

3.2. Metodologia

Numa perspectiva teórica, a metodologia adoptada segue os princípios dos Modelos Organizacional e do Ciclo de Vida em Cascata do Projecto.

O Modelo Organizacional caracteriza-se por fazer uso da matemática, da estatística e da economia, combinado com as ciências sociais. [10] Nesta metodologia, o suporte à decisão de negócio tem como critério a obtenção de um resultado final satisfatório para o cliente. Este modelo faz estimativas e considera a existência de limitações ao nível dos recursos humanos, tempo e custos. O processo de decisão é geralmente tomado a curto prazo e muito orientado para os resultados finais, possuindo um carácter estratégico, na perspectiva de negócio.

O Modelo em Cascata estabelece um faseamento de tarefas preestabelecido, durante o ciclo de vida de um projecto.

3.2.1. Metodologia do ciclo de vida do Projecto

A metodologia seguida no Departamento é bem definida e visa garantir a qualidade do produto, durante o ciclo de vida do projecto.

Nos primórdios da Engenharia do Software o Modelo em Cascata (ou *Waterfall*) era o mais utilizado durante o ciclo de vida de um projecto. Muitos dos modelos usados nos nossos dias possuem similaridades ou derivam do Modelo em Cascata. Este caracteriza-se por possuir diversas fases, bem definidas, em que a próxima fase só é iniciada se a imediatamente anterior estiver totalmente concluída. Esta característica pode revelar-se uma falha dada a sua inflexibilidade. O carácter dinâmico de uma área de negócios obriga a uma constante revisão dos requisitos. Por outro lado, a satisfação do cliente é uma prioridade, como tal o carácter humano dos clientes determina algumas alterações inesperadas que originam retrocessos nas diversas fases do ciclo de vida de um projecto.

O Modelo Organizacional está também embebido na metodologia seguida neste projecto, pois foram definidas regras que usam cálculos matemáticos e estatísticos, patenteando uma avaliação criteriosa dos resultados obtidos.

Resumindo, ao longo do ciclo de vida de um projecto é necessário seguir uma metodologia algo flexível em que existe a noção de interacção entre os responsáveis pelo desenvolvimento do projecto e os clientes. Todos estes motivos condicionam a aplicação rígida do Modelo em Cascata, originando modelos alternativos que se concretizam numa adaptação e reestruturação do modelo em Cascata original.

Numa panorâmica esquemática, as equipas do Departamento seguem uma metodologia orientada por fases bem definidas, tal como se representa abaixo. [Figura 11]

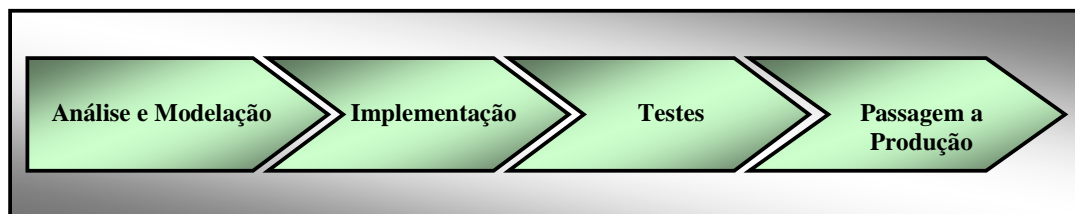


Figura 11 – Fases da Metodologia de Projecto usada no Departamento

Esta metodologia é adaptada à natureza de cada projecto. Para projectos de maior complexidade, esta metodologia pode ser expandida de modo a considerar um faseamento mais detalhado. [Figura 12]

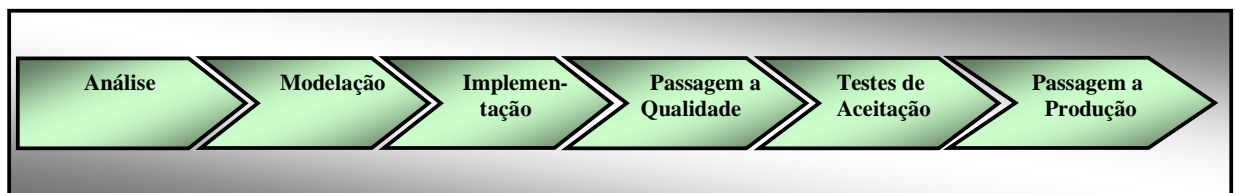


Figura 12 – Detalhe das Fases da Metodologia de Projecto usada no Departamento

Durante o ciclo de vida deste projecto, cada uma das fases compreendidas na metodologia estão bem definidas, com vista a cumprir todos os objectivos e solucionar os problemas existentes com sucesso.

Em seguida caracteriza-se, com detalhe, cada uma das fases metodológicas anteriormente referidas. [Tabela 5]

Fases da Metodologia	Definição
Análise	Fase inicial que consiste no levantamento de requisitos e na análise dos objectivos, com vista à resolução do problema. Numa perspectiva de negócio, primeiro definem-se todas as regras a cumprir para, no final, se garantir a melhor solução.
Modelação (ou Desenho)	Fase em que se efectuam os modelos arquitectural e lógico, de acordo com a estrutura de dados existente no projecto. Assim, efectua-se o desenho de acordo com a resolução do problema, proposta na fase anterior.
Desenvolvimento (Implementação)	Fase caracterizada pela implementação/desenvolvimento, em concordância com as soluções das fases de análise e modelação. No âmbito deste projecto, este passo inclui todo o ambiente de <i>Reporting</i> e processos de exploração, mapeamento e listagem de dados. Antes de se passar à fase seguinte são feitos alguns testes que visam asseverar o cumprimento dos objectivos inicialmente propostos e o bom funcionamento da solução implementada.
Passagem a Qualificação (Qualidade)	Fase definida pela simulação da passagem a produção, na perspectiva do cliente. Assim, consiste numa tentativa de encontrar algum incumprimento que comprometa o funcionamento correcto da solução, caso esta já se encontrasse em ambiente de produção. Tendencialmente, este passo é executado prevendo a óptica final do projecto, sempre de acordo com a perspectiva do cliente/utilizador final do sistema.
Testes de Aceitação	Nos mesmos moldes da fase anterior, existem colaboradores internos ou externos e/ou utilizadores finais do projecto que efectuam testes de aceitação, comprovam a validade da solução e das funcionalidades do projecto, detectando possíveis erros existentes nos dados. Neste projecto, o cliente integra a fase de testes. Consoante a natureza do projecto, são utilizadas diversas técnicas de teste.
Passagem a Produção	Fase em que o projecto é remetido para o ambiente de produção. É entregue ao cliente como solução final de resposta ao problema inicialmente existente. Quaisquer alterações necessárias ficam a cargo da equipa de manutenção. Deverá existir documentação técnica que possa ser consultada como manual de utilização do ambiente de produção.

Tabela 5 – Definição de cada uma das Fases da Metodologia

Como em todas as metodologias, a natureza e as necessidades do projecto condicionam a aplicação de cada uma das fases anteriormente descritas.

No caso específico do projecto desenvolvido, as fases anteriormente caracterizadas são seguidas à medida que as necessidades vão emergindo.

Face a outros projectos existentes nesta instituição, os motivos e que fazem com que a metodologia seja seguida de modo flexível residem no facto deste projecto ser:

- De menores dimensões face aos projectos bancários existentes, facto que deriva do tempo reduzido do projecto – 9 meses;
- Um projecto constituído por um só colaborador e pela interacção com os clientes, não estando afecta uma equipa de desenvolvimento.

Dada a natureza do projecto, esta metodologia não foi aplicada de forma rígida. O faseamento metodológico foi seguido, de modo flexível, à medida que as necessidades iam surgindo e a par do cumprimento dos objectivos e prazos de resolução do problema existente.

3.2.2. Metodologia Adoptada – Detalhe

A metodologia adoptada durante o ciclo de vida deste projecto consiste numa adaptação da metodologia anteriormente descrita, de modo a cumprir as soluções e os objectivos anteriormente mencionados, na perspectiva de alcançar ganhos precisos e consistentes a diversos níveis. Usaram-se também conceitos dos modelos Organizacional e em Cascata.

Com maior pormenor, a metodologia seguida edificou-se nos passos que se passam a referir:

- 1. Análise e definição precisa dos objectivos do projecto:**
 - 1.1.** Especificação dos requisitos.
- 2. Desenho e Análise Preliminar do Sistema:**
 - 2.1.** Exploração da Informação [TOAD e MSTR]:
 - 2.1.1.** Compreensão das Estruturas de Dados em PL/SQL;
 - 2.1.2.** Categorização do tipo de objectos existentes;
 - 2.1.3.** Identificação de dependências;
 - 2.2.** Definição do âmbito das regras.
- 3. Desenvolvimento/Implementação da componente de análise e gestão:**
 - 3.1.** Definição precisa de regras: métricas, filtros e funções a aplicar nos relatórios de gestão administrativa implementados;
 - 3.2.** Construção de Relatórios;
 - 3.3.** Comparação de Resultados intermédios obtidos;
 - 3.4.** Categorização de objectos de acordo com a sua periodicidade:
 - 3.4.1.** Agregação de Informação proveniente de camadas lógicas distintas.
- 4. Produção de Resultados de Gestão de Metadados:**
 - 4.1.** Análise de Resultados;
 - 4.2.** Estruturação dos Resultados Produzidos:
 - 4.2.1.** Produção de listagens com todos os resultados obtidos;
 - 4.2.2.** Produção de resultados em diferentes formatos, para melhor visualização.
 - 4.3.** Reanálise e comparação de resultados obtidos usando um método alternativo;
 - 4.4.** Produção de documentação;

- 5. Passagem a Qualificação e Metodologia modular:**
 - 5.1.** Envio de resultados para utilizadores, para análise de pertinência dos resultados obtidos;
 - 5.2.** Espera de feedback do cliente;
 - 5.3.** Recepção da segunda análise;
 - 5.4.** Comparação de Resultados obtidos vs Resultados enviados pelos utilizadores;
 - 5.5.** Adequação da solução;
 - 5.6.** Produção de documentação.
- 6. Implementação de alterações e Testes de Aceitação:**
 - 6.1.** Remoção dos objectos obsoletos indicados pelos utilizadores;
 - 6.2.** Início de fase de indisponibilização: fase intermédia anterior à remoção definitiva em que os utilizadores poderão indicar a falta de algum objecto proposto para remoção: fase de testes;
 - 6.3.** Análise de Impactos;
 - 6.4.** Fase de remoção efectiva de objectos obsoletos;
 - 6.5.** Produção de documentação.
- 7. Integração do Sistema de *Reporting* de Gestão desenvolvido no Sistema de *Reporting* existente:**
 - 7.1.** Passagem a Produção;
 - 7.2.** Avaliação de impactos e de resultados obtidos.
- 8. Manutenção do sistema.**
- 9. Utilização do trabalho desenvolvido neste projecto, noutros existentes.**

A metodologia referida nesta subsecção será alvo de maior pormenor no próximo capítulo. Por vezes, foram efectuadas fracções deste faseamento em paralelo. Melhorias, comparações e documentação intermédia foram desenvolvidas de acordo com as necessidades existentes. Assim, conseguiu-se uma metodologia flexível e orientada a objectivos específicos, dependentes da fase que estava em curso.

Anteriormente foi referido o Modelo em Cascata no âmbito de projectos de Engenharia do Software, em que cada passo é concluído antes que o próximo passo possa ser aberto. Por exemplo, a especificação de requisitos é iniciada e tem de ser terminada antes que se iniciem as fases de a análise e desenho do sistema. A nomenclatura e definição de cada passo pode variar mas o ciclo de vida começa com os requisitos seguindo-se a fase de desenho, implementação, testes e por fim a manutenção do sistema. Esta sequência lógica do Modelo em Cascata foi, nesta medida, adoptada neste projecto. Em conjunto com a metodologia anteriormente descrita. Deste modo, a limitação da estrutura pouco flexível do Modelo em Cascata foi ultrapassada, dado houve “espaço de manobra” para a revisão de passos iniciais em fases mais tardias do ciclo de vida do projecto. Assim, está garantida a interacção com o cliente, a adaptação da solução às necessidades existentes e a flexibilidade da metodologia do ciclo de vida do projecto, ao longo do tempo. O Modelo Organizacional adoptado baseia-se nas medidas matemáticas, no uso da estatística e na avaliação quantitativa e qualitativa dos resultados obtidos.

3.3. Planeamento

Após a fase de aprovação de um projecto, a empresa de acolhimento externa atribui os recursos necessários e define, com rigor, as tarefas a estes associadas.

Numa aplicação própria para o efeito, todos os meses, são imputadas as horas dispendidas numa determinada tarefa. Este processo de introdução de horas é efectuado diariamente. Deste modo, os recursos têm a noção do tempo consumido em cada tarefa e podem fazer estimativas do seu rendimento ao longo do tempo.

Este método de trabalho obriga a uma constante preocupação em cumprir prazos predefinidos para a alocação de recursos/tempo, melhorando a performance e o sentido de responsabilidade dos colaboradores.

Nesta secção apresenta-se o planeamento inicial e o planeamento efectivamente concretizado, mencionando-se os motivos de desvios existentes.

3.3.1. Plano de Tarefas Inicial

A metodologia seguida, apresentada na secção anterior, foi segmentada em fases distintas. No entanto, este projecto foi inicialmente segmentado em quatro fases principais, de acordo com uma base temporal, tal como foram definidos no arranque deste estágio. [Tabela 6]

Setembro (Fase 1) – Enquadramento no Projecto
▪ Introdução à Ferramenta MicroStrategy;
▪ Aprendizagem de projectos implementados em Produção;
▪ Introdução ao Enterprise Manager (Componente de Administração MSTR);
▪ Modelo de Dados da Data Warehouse.
Outubro a Dezembro (Fase 2) – Levantamento e Análise Metadata Actual
▪ Levantamento integral dos objectos MicroStrategy existentes: atributos, métricas, factos, filtros, prompts, tabelas, entre outros;
▪ Análise detalhada de todos os objectos dependentes;
▪ Identificação de impactos;
▪ Avaliação de redundâncias;
▪ Escrita e Entrega do Relatório Preliminar.
Janeiro a Abril (Fase 3) – Proposta de Optimização e Definição Metadata Objectivo
▪ Identificação de Objectos a excluir ou sistematizar e todas as dependências implicadas;
▪ Apresentação de optimizações;
▪ Desenho detalhado da Metadata Objectivo;
▪ Produção de Dossier de Especificações – Levantamento e Metadata Objectivo.
Maio (Fase 4) – Realização de Exemplo de Optimização
▪ Selecção de exemplo prático para execução;
▪ Implementação de optimização;
▪ Testes de não regressão;
▪ Identificação/caracterização dos ganhos atingidos.

Tabela 6 – Planeamento Inicial de Tarefas Segmentado em Quatro Fases Principais

Em seguida, apresenta-se, de modo resumido, o planeamento inicialmente proposto, no formato de Mapa de Gantt. [Figura 13]

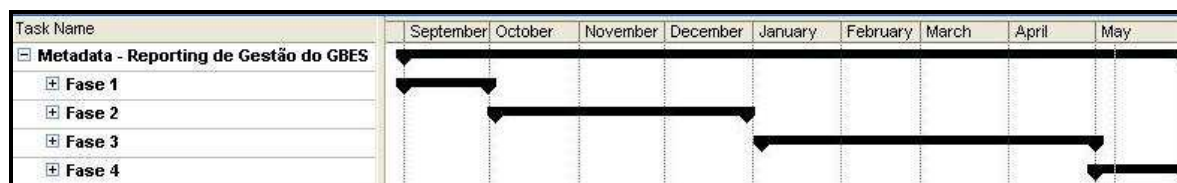


Figura 13 – Mapa Gantt do Planeamento inicialmente definido

3.3.1.1. Plano de Tarefas Efectuado

Dado o nível de detalhe deste planeamento, por questões que se prendem com a dimensão e apresentação este encontra-se colocado em anexo. [Anexo 3]

3.3.1.2. Plano de Tarefas Inicial versus Concretizado

O planeamento de tarefas inicialmente proposto foi cumprido na sua totalidade. Algumas subtarefas e algum reajustamento de objectivos foram sendo adicionados a cada uma das fases existentes, concretizando-se num plano de tarefas realizado com maior detalhe e aplicado a uma metodologia de ciclo de vida do projecto bem definida.

Em relação a desvios existentes face ao planeamento inicialmente proposto, existiu uma antecipação na concretização de algumas fases do plano. Esta antecipação resultou de um esforço constante na realização de tarefas e numa urgência na produção de resultados, por parte do cliente.

O adiantamento resultante da concretização deste projecto, propiciou as condições necessárias ao início de um outro que dá continuidade a todo este trabalho desenvolvido.

Este trabalho abarca um conjunto de conceitos e critérios que servem de ponto de partida para um projecto complexo e de maiores dimensões, a realizar de futuro.

Através da análise dos dois planos, sobressaem duas visões. A do Mapa Gantt do Planeamento inicialmente definido, está orientada para uma perspectiva de concretização de tarefas, numa perspectiva temporal mensal. O Planeamento Detalhado do Mapa de Gantt, em anexo, está orientado para uma perspectiva metodológica em fases, de acordo com a metodologia seguida durante o ciclo de vida do Projecto. O plano detalhado reflecte os conceitos da metodologia adoptada, de acordo com a que é seguida no departamento. Esta usa noções do modelo tradicional em Cascata e adapta-se às necessidades do Projecto, de modo a cumprir os objectivos para este delineados.

Sumário

Neste capítulo foram descritos os objectivos a atingir, após a identificação dos problemas e necessidades existentes. Foi feita uma descrição da solução proposta a diversos níveis. A metodologia adoptada de modo a cumprir a totalidade dos objectivos, ao longo do ciclo de vida do projecto foi também alvo de dissertação. Finalmente, foi apresentado o planeamento inicial de tarefas e a comparação com o plano efectivamente concretizado.

Capítulo 4

Trabalho Realizado

Neste capítulo apresenta-se todo o trabalho concretizado. É feita uma apresentação de todas as ferramentas e tecnologias utilizadas. Faz-se uma descrição pormenorizada dos resultados obtidos e da sua avaliação, com base nas medidas de precisão e abrangência dos mesmos. Faz-se uma análise crítica das ferramentas e tecnologias subjacentes ao projecto. Finalmente, apresenta-se as contribuições deste projecto e o enquadramento destas num trabalho a realizar de futuro.

No contexto deste projecto, existem condições iniciais que dão início ao trabalho que foi sendo desenvolvido. Com o propósito de colmatar problemas inicialmente existentes, aplicou-se, durante o ciclo de vida do projecto, uma solução a diversos níveis e foi seguida uma metodologia definida de acordo com as necessidades do projecto, num processo de Engenharia.

A crescente necessidade de gestão da informação e a grande quantidade de dados existentes criaram as condições que deram origem ao Sistema de *Reporting* de gestão administrativa, desenvolvido neste projecto.

Os requisitos iniciais a satisfazer consistiam:

- No cumprimento das tarefas dentro dos prazos predefinidos;
- Na análise da totalidade dos dados presentes no repositório de metadados;
- Na aplicação de uma metodologia modular que inclui a participação do cliente/utilizador final no processo de avaliação dos resultados intermédios obtidos;
- Na remoção efectiva da totalidade de objectos obsoletos, após a análise de pertinência do cliente/utilizador;
- Em ganhos em termos de memória física ocupada e consequente diminuição de carga do Sistema;
- Na obtenção de um ambiente de Reporting organizado, automático, reutilizável e eficiente, pronto a ser usado no âmbito de projectos futuros.

No decurso deste projecto foram cumpridas as tarefas estipuladas no planeamento, sendo considerada uma alocação de tempo a outras tarefas, que consistiram:

- Na familiarização com novas ferramentas;
- Exploração e aprendizagem da estrutura de dados;
- Leitura de documentação técnica.

O uso de ferramentas específicas complementa o trabalho desenvolvido neste projecto, sendo impreterível a sua utilização.

Foram empregadas técnicas de pós-processamento de informação, bem como análises de pertinência de objectos, testes de não regressão e técnicas de avaliação dos resultados obtidos.

4.1. Responsabilidades no Projecto

Neste projecto, não existiu uma equipa afecta motivo pelo qual acompanhei a totalidade do mesmo, durante o seu ciclo de vida.

A natureza do projecto faz com que não tivesse sido desenvolvido de raiz, uma vez que todos os dados já se encontravam em produção. Tratou-se de uma integração num processo contínuo de gestão do repositório de metadados. No entanto, este projecto caracteriza-se pela aplicação deste processo de gestão do conhecimento de modo mais intenso e com objectivos bem definidos, que se traduzem na concretização de um ambiente de *Reporting* de gestão administrativa organizado, reutilizável e de qualidade. A diminuição de carga do Sistema global de *Reporting* tem igualmente repercussões benéficas para o mesmo.

O âmbito deste projecto cruza diversos níveis de agrupamento da informação, passando pelos utilizadores finais/clientes e fazendo uso de um conjunto de ferramentas e tecnologias.

Em seguida faz-se um resumo esquemático introdutório das principais tarefas, ferramentas e do fluxo arquitectural de informação que se irão abordar no presente capítulo. [Figura 14]

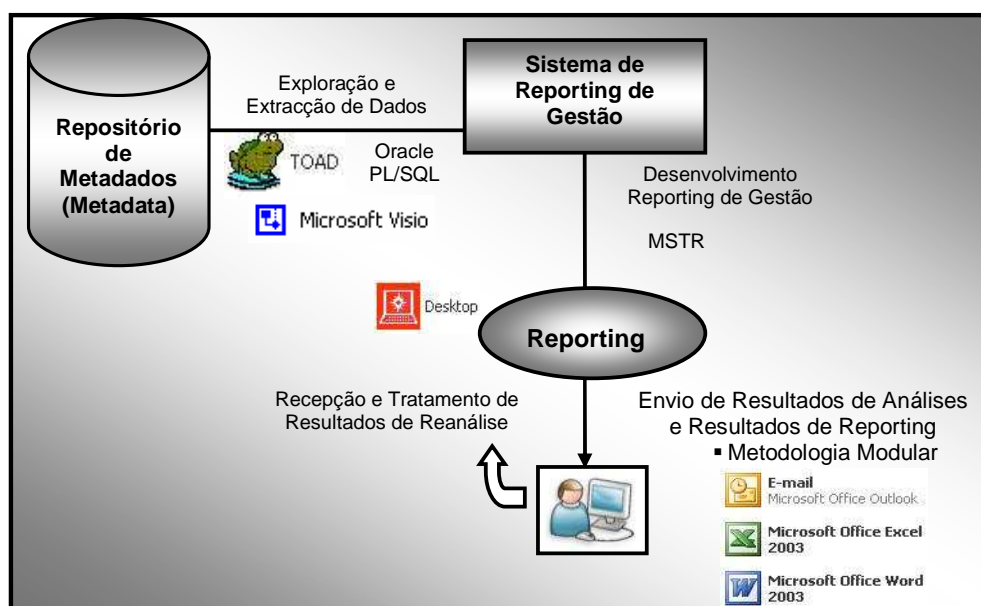


Figura 14 – Resumo de Tarefas e Ferramentas

4.2. Ferramentas Utilizadas

No âmbito deste projecto, foram usadas diversas ferramentas que se passam a descrever em seguida.

4.2.1. MicroStrategy

O MicroStrategy consiste numa poderosa ferramenta de *Reporting*, amplamente usada em ambientes empresariais, nomeadamente as de carácter bancário. Permite desenvolver relatórios (*reports*) integrados numa mesma metadata. Possibilita uma análise contínua e a monitorização de performance de negócio fornecendo, aos utilizadores finais, as fontes de informação que se revelem necessárias para melhorar as decisões de negócio.

Os relatórios podem ser disponibilizados via Web ou através do Desktop. Os utilizadores com maior experiência usam o Desktop, sendo esta uma vertente mais robusta que, para além de apresentar os reports existentes, permite que estes efectuem *query ad-hoc*, criação de filtros, métricas e *prompts* sobre os dados existentes.

Sob o ponto de vista de desenvolvimento a criação de projectos, nesta ferramenta, consiste no ponto mais alto de intersecção entre um DW, um repositório de Metadados e os clientes/utilizadores finais do Sistema de *Reporting*. Conceptualmente, um projecto, consiste no ambiente em que todo o *Reporting* relacionado é efectuado. Um projecto típico contém a totalidade dos objectos, sendo estes filtros, métricas, prompts, atributos, tabelas, etc. Quando agregados, num mesmo relatório, resultam na disponibilização da informação de negócio pretendida, num formato intuitivo, exportável e organizado – o relatório ou *report*.

De modo resumido, apresentam-se alguns dos objectos existentes no Repositório de Metadata, no âmbito da ferramenta de *Reporting* MicroStrategy.

[Tabela 7]

Conceito	Definição
Atributo	Representação e agrupamento de categorias e de níveis de dados. Não são numéricos e representam informação por assunto para caracterizar e qualificar um grupo de dados. Respondem a perguntas sobre um facto.
Facto	Conceptualmente é visto como uma medida de registo das ocorrências de negócio. Factos consistem numa situação única, dada uma certa combinação de atributos. Permitem o acesso a dados armazenados no DW; consistem na base dos relatórios e na maioria das análises requeridas pelos utilizadores. Na sua maioria são valores numéricos.
Filtro	Define condições lógicas que limitam o intervalo de dados numéricos apresentados num relatório. Limita os valores numéricos de um dado atributo a um intervalo de valores.
Métrica	Consiste num cálculo analítico obtido através de uma expressão construída a partir de objectos existentes, sendo estes: funções, atributos, factos ou outras métricas.
Relatórios (Reports)	Consistem numa síntese e integração de dados cujos resultados servem um vasto conjunto de propósitos. Visualização de informação para ajudar o suporte à decisão e visão de negócio. Podem ser administrativos, ie, resultam da análise dos dados como processo de gestão de informação interna sobre a qualidade dos dados. Consistem numa <i>query</i> SQL cujos resultados são apresentados numa forma gráfica, <i>user-friendly</i> através das ferramentas de <i>Reporting</i> .
Tabela	Na perspectiva de Sistemas de Informação, consiste no componente físico primário de um DW / Base de Dados, constituído por colunas e dados dos mais variados tipos.

Tabela 7 – Definição de Conceitos MicroStrategy

O MicroStrategy é constituído por diversos componentes. O componente mais usado no decurso deste projecto foi o Enterprise Manager. Consiste no componente administrativo desta ferramenta de *Reporting*, onde é possível o desenho e implementação de relatórios administrativos de gestão em modo gráfico, visualização e edição do código SQL subjacente a um relatório, definição de regras (métricas, filtros e *prompts*). Possibilita a produção de resultados com informação estatística. Todo este ambiente de *Reporting* administrativo decorrente do Enterprise Guide, fornece um modo automático de reutilização dos relatórios para análises e processos de gestão de informação futuras.

Em seguida faz-se uma panorâmica do ambiente MicroStrategy, através da descrição e apresentação visual dos seus componentes.

O Desktop da ferramenta MicroStrategy está representado na figura abaixo. [Figura 16]

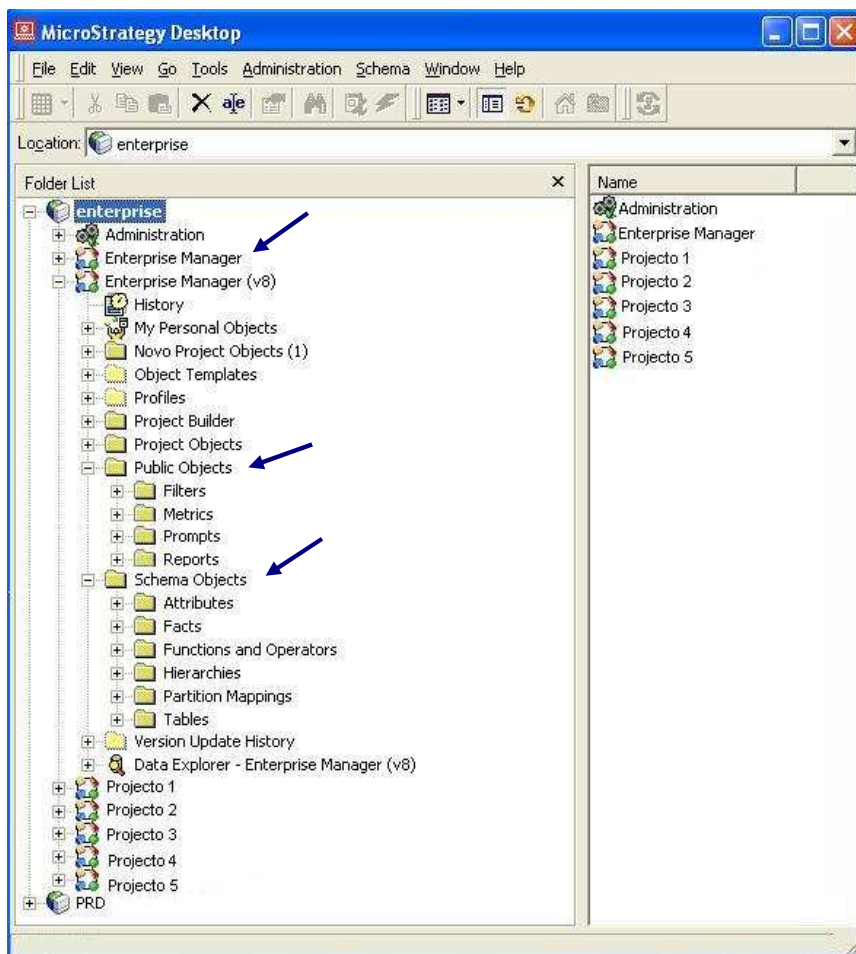


Figura 15 – Ambiente MicroStrategy Desktop (*)

O MicroStrategy, vertente Desktop, possibilita o acesso ao Enterprise Manager e a diversos tipos de objectos.

Esta ferramenta define dois níveis de organização de objectos. Um nível consiste num modelo lógico e outro está a um nível físico, sendo estes:

- **Public Objects**: consistem num modelo lógico de dados, caracterizado por ser uma “fotografia” de toda a estrutura dos dados e como estes se relacionam num ambiente de negócio, técnico ou conceptual. Contém relatórios, filtros, métricas, entre outros.
- **Schema Objects**: consistem num modelo físico, baseado no modelo lógico, organizado de modo a que este possa ser visto numa perspectiva de Bases de Dados, contendo factos, atributos, tabelas, etc.

(*) Por motivos que se prendem com dados confidenciais, os projectos que representam diferentes níveis de agrupamento de informação, foram denominados genericamente por Projecto1, Projecto 2, ..., Projecto 5.

Para além das características referidas para o Desktop, o MicroStrategy possui um componente de implementação de relatórios, mostrado na próxima figura. [Figura 17]

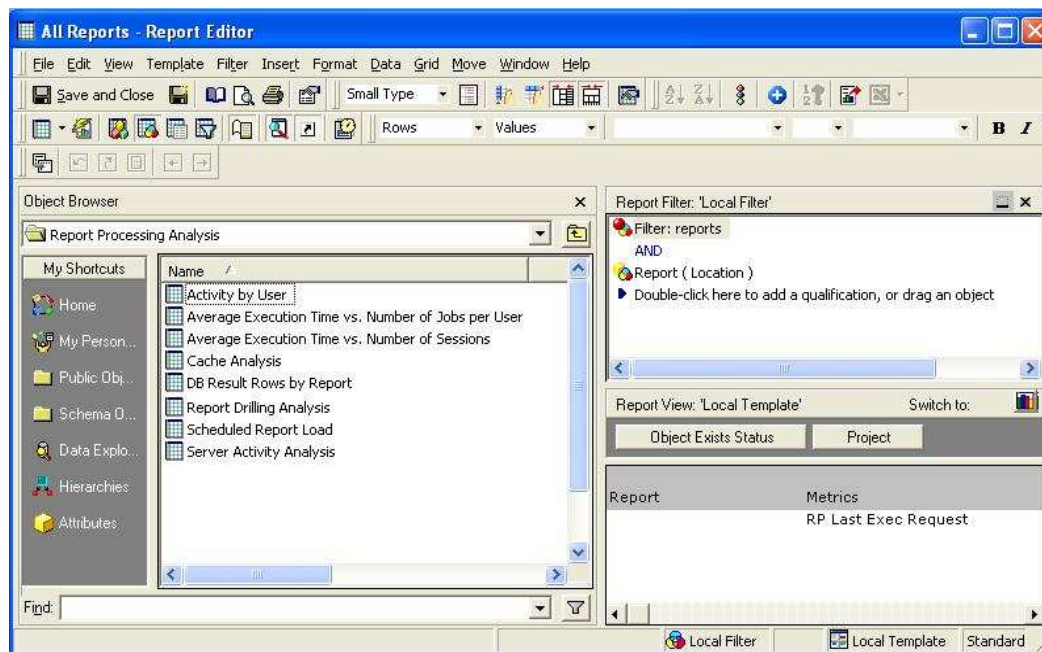


Figura 16 – Ambiente de Criação de Relatórios

Este ambiente de construção de relatórios permite funcionalidades de *drag-and-drop* na escolha de componentes do relatório. Permite escolher o modo de visualização do relatório, a definição de métricas, *prompts* e filtros, entre muitas outras funcionalidades. É também possível visualizar e editar o código PL/SQL subjacente a um relatório.

Os relatórios construídos permitem filtrar informação estatística, produzir relatórios com informação de negócio, consistindo no ambiente típico de *Reporting* de auxílio à tomada de decisões, no âmbito de uma organização. Um conjunto de relatórios relacionados está agrupado, nesta ferramenta, por projecto representando diferentes dimensões de negócio.

No MicroStrategy, um projecto consiste no nível mais alto de intersecção entre um DW, um repositório de metadados e os utilizadores/clientes MicroStrategy. Tipicamente, sempre que se inicia uma sessão no MicroStrategy está a aceder-se a um projecto no qual estão definidos os objectos MicroStrategy correspondentes.

O MicroStrategy possui uma Base de Conhecimento (*Knowledge Base*) que contém as mais variadas informações, acerca da ferramenta. No âmbito deste projecto, utilizou-se um conjunto de informação de natureza estatística que permite prever o comportamento de um sistema. Este prognóstico permite uma análise crítica de resultados, sob as perspectivas de previsto/estimado versus concretizado.

4.2.2. TOAD

O TOAD consiste numa ferramenta que possibilita o acesso aos dados e o desenvolvimento de código na linguagem PL/SQL em Bases de Dados Oracle.

Esta ferramenta fornece um ambiente gráfico bastante intuitivo, em janelas. Contém uma janela para edição e pesquisa em SQL. Tem um componente de navegação que possibilita a visualização e administração da totalidade de tabelas e *views* existentes numa Base de Dados Oracle.

No decurso deste projecto o TOAD permitiu:

- A análise e pesquisa de dados presentes no repositório de metadados;
- Exploração e familiarização com as estruturas de dados existentes;
- Percepcionar a estrutura física e desenho dos dados;
- Encontrar relações e dependências existentes;
- Comparar os resultados obtidos, por programação em PL/SQL, aquando do uso de outras ferramentas;
- Filtrar a informação necessária e combiná-la com outras ferramentas;
- Extrair resultados de *queries* e exporta-los para ficheiros em diversos formatos (ex. Excel, txt, xml).

Todas as funcionalidades mencionadas fizeram do TOAD [Figura 17] uma ferramenta de uso contínuo e da maior importância no decurso deste projecto.



Figura 17 – Ambiente de Desenvolvimento TOAD

Em anexo, apresentam-se, em maior detalhe, os componentes e funcionalidades desta ferramenta. [Anexo 1]

4.2.3. Folhas de Cálculo

Neste projecto foram usadas com muita frequência folhas de cálculo do Microsoft Excel 2003. Revelam-se uma boa base de estruturação e apresentação de resultados de análise obtidos, no âmbito de *Reporting* e filtragem de objectos obsoletos. As ferramentas MicroStrategy e TOAD possuem funcionalidades de exportação de resultados de análises e relatórios para formato de folhas de cálculo.

Embora a visualização de relatórios possa ser efectuada via Web ou integrada na ferramenta MicroStrategy, no passado as folhas de cálculo eram o único meio de disponibilização dos mesmos. No entanto, com o passar do tempo, estas continuam a ser um meio eficaz de visualização, armazenamento, análise e edição de resultados de relatórios para o utilizador final.

Neste projecto, todos os levantamentos analíticos de informação, relativos a objectos, foram exportados para folhas de cálculo Excel e enviados ao cliente.

Foram usados componentes do Excel que proporcionam um ambiente de visualização e análise de informação mais organizado, de entre estes destacam-se os *Pivot Tables* e *Pivot Charts* (gráficos).

- *Pivot Charts*: consistem em gráficos que proporcionam um ambiente visual de comparação e compilação dos resultados de análises obtidos. A apresentação gráfica em conjunto com a disponibilização de totais (*counts* e *sums*) proporciona uma visualização imediata e intuitiva dos resultados.
- *Pivot Tables*: são um formato de representação de dados direccionados para a análise, contabilização, cruzamento e representação sumária de todos os relatórios existentes.

A função *VLOOKUP* do Excel foi muito usada do âmbito deste trabalho pois possibilita a comparação de campos, facilitando a análise de dados de tabelas em paralelo.

4.2.4. Processamento de Texto

As ferramentas de processamento de texto são muito utilizadas aquando da produção de documentação técnica. No âmbito deste projecto foi usado o Microsoft Word 2003.

4.2.5. Visio

O Microsoft Visio foi usado na construção de modelos de mapeamento de estruturas de dados, arquitectura de informação, entre outros.

É uma ferramenta intuitiva que permite uma fácil construção de modelos e esquemas dos mais diversos tipos.

4.2.6. Microsoft Project

Para efectuar o planeamento de tarefas, no formato de Mapas de Gantt, foi utilizado o Microsoft Project. Esta ferramenta é muito útil aquando do escalonamento de recursos a projectos e respectivas tarefas. Com base nesta ferramenta, consegue-se atingir uma metodologia de trabalho organizada e com objectivos e metas temporais predefinidas a cumprir.

4.3. Detalhe das Tarefas do Trabalho Realizado

Todas as tarefas concretizadas estão de acordo com o planeamento de trabalho e a metodologia anteriormente descrita.

Neste projecto as tarefas de maior enfoque consistiram na análise e gestão da informação presente no repositório de metadados, bem como o desenvolvimento de *Reporting* administrativo de gestão.

A tarefa de análise de informação examina mais do que o componente tecnológico, não podendo, neste caso de estudo, ser um processo completamente automático dada a natureza da informação em questão. A desagregação de alguns níveis de informação não permite a exploração directa dos dados, havendo o recurso ao uso sistemático de programação – *queries* PL/SQL – de modo a filtrar o detalhe pretendido.

A gestão de informação consiste no processo de converter informação em acção e naquilo que pode ser feito com a totalidade do conhecimento.

Foram usadas técnicas de pós-processamento de dados resultantes de todo o processo de solução aos problemas inicialmente identificados, com o intuito de tratar a informação.

O desenvolvimento de relatórios administrativos de gestão foi efectuado com o componente Enterprise Manager da ferramenta de *Reporting* MicroStrategy, complementado com desenvolvimento de programação em PL/SQL na ferramenta TOAD.

Em seguida, particulariza-se o trabalho concretizado em todas as fases da metodologia adoptada, descrita no capítulo anterior.

4.3.1. Análise, Especificação de Requisitos e Modelação

Esta fase dá início à metodologia adoptada durante o ciclo de vida do projecto. Foram definidos os objectivos base para a resolução do problema existente, o que se traduz, em termos latos, na análise, especificação de requisitos e definição de modelos informais.

A natureza e o âmbito do projecto foram estudados detalhadamente com o propósito de apreender os conceitos técnicos, arquitecturais e de negócio. Foram criados modelos informais que caracterizam a estrutura da informação, permitindo a apreensão da arquitectura física e lógica subjacente ao ambiente DW em questão.

Nesta fase foi também efectuado o desenho preliminar do Sistema, como modo de identificar objectivos precisos e trabalho inicial a realizar. [Figura 18] Nesta tarefa foi feita uma exploração intensiva da informação, usando as ferramentas MicroStrategy, TOAD, Excel e programação em PL/SQL. Foi realizado, num processo de análise, o levantamento categórico da totalidade dos objectos MicroStrategy existentes, bem como a identificação das dependências inter-objectos.

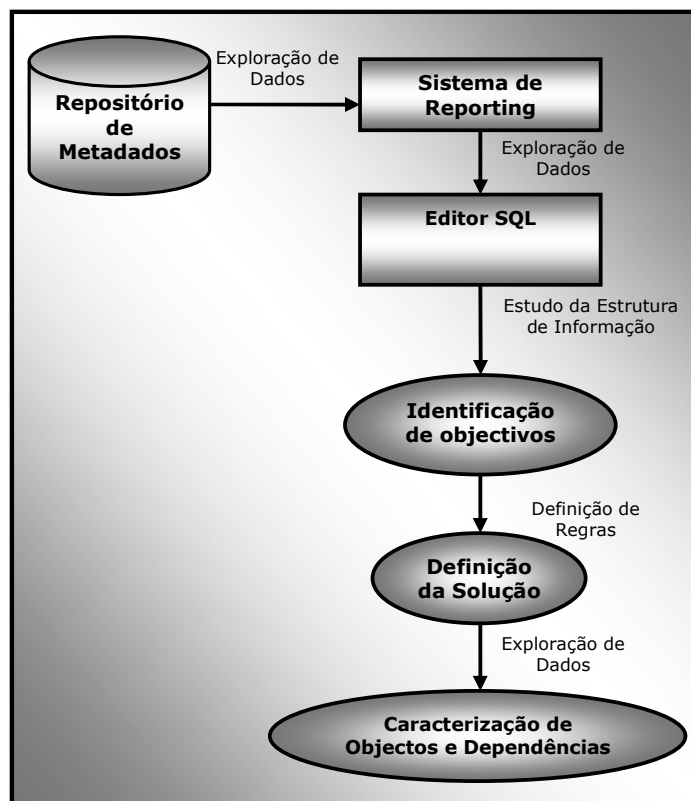


Figura 18 – Desenho preliminar decorrente da fase de análise

Depois da definição da solução, a diversos níveis, foi efectuada uma reunião formal com o cliente para revisão e aceitação das regras a aplicar, de modo a validar os objectivos a alcançar e o *range* das regras.

Em termos de desenho/modelação a exploração de informação permitiu identificar o Modelo DW Arquitectural [Figura 19] e o de Modelação Dimensional existentes.

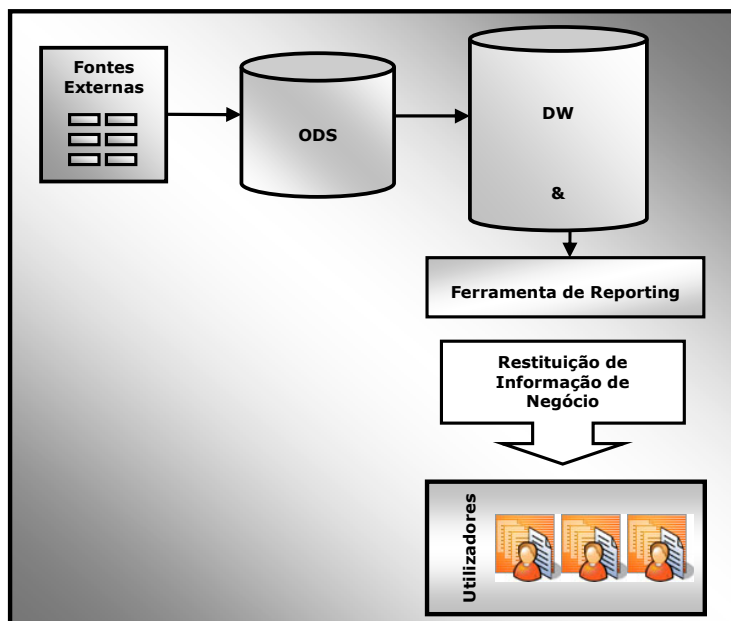


Figura 19 - Modelo Arquitectural DW

O Modelo Dimensional existente é na forma de Esquema em Estrela, em que existem tabelas centrais de factos ligadas a tabelas de dimensões, concretizando-se numa modelação lógica das estruturas de dados.

Esta fase permitiu também identificar os objectos mais pertinentes no repositório de metadados. Os objectos, ao nível do modelo dimensional, identificados como sendo de maior relevância são:

- **Relatórios**

Os relatórios encontram-se agrupados por áreas de negócio e produtos, estando catalogados em diferentes projectos, tanto no repositório de Metadados como na ferramenta MicroStrategy. Cada um destes projectos representa âmbitos bancários distintos. Contudo, ao nível da informação, todos se estruturam na mesma lógica funcional.

Numa perspectiva de modelação dimensional estes estão definidos em tabelas, factual de relatórios e dimensões. São descritos por um conjunto de registos que os caracterizam e definem a sua localização física de armazenamento.

Representam o nível mais próximo de contacto com o cliente/utilizador, uma vez que a restituição da informação se faz via relatório.

- **Utilizadores**

Representam utilizadores das diversas áreas bancárias correspondendo a perfis de utilização distintos.

Estão definidos em tabelas, factual de utilizadores e dimensões, no modelo dimensional lógico.

Possuem um conjunto de atributos que os caracterizam.

Estão associados a uma tabela de sessões, permitindo identificar datas de sessão/login de cada utilizador.

- **Tabelas**

No repositório, estão mapeadas como tendo sempre uma *View* associada a cada tabela. As tabelas possuem diferentes âmbitos, dependendo da secção de informação a que pertencem.

Para cada um destes objectos foram estabelecidas regras que permitiram rastrear desuso ou desactualização.

A definição de regras de filtragem de informação foi decidida em duas vertentes, aquando da análise das características dos registos de cada uma das tabelas de objectos (factual e dimensões) em conjunto com o cliente. Estas variam consoante o tipo de objecto, sendo estas: [Tabela 8]

Objecto	Regras
Relatórios	<p>Regra Temporal: Considera-se um relatório como sendo obsoleto se não tem utilização posterior a 2005. <u>Regra:</u> Data da Última Execução <= '2005-01-01'</p> <p>Regra de Pertença: Relatórios de teste podem ser removidos de acordo com o âmbito a que pertencem, salvo indicação do <i>owner</i>.</p> <p>Regra de Duplicação: Relatórios em duplicado podem ser removidos.</p>
Utilizadores	<p>Regra Temporal: Considera-se um utilizador como sendo obsoleto se não tem utilização posterior a 2005. <u>Regra:</u> Data da Última Sessão <= '2005-01-01'</p>
Tabelas	<p>Regra de Característica: Como cada tabela tem uma <i>view</i> a esta associada, todas as tabelas sem <i>views</i> que lhes correspondam poderão ser removidas.</p> <p>Regra de Utilização: Cada tabela possui um registo que indica o tamanho, tabelas com tamanho a zero indicam uma não-utilização.</p>

Tabela 8 – Regras definidas para cada tipo de Objecto

Esta primeira fase de trabalho foi de extrema importância pois permitiu a familiarização com ferramentas, tecnologia, apreensão de teoria sobre DW e arranjo arquitecturas das estruturas de dados, conceitos fulcrais para a concretização, com sucesso, das fases seguintes.

4.3.2. Desenvolvimento/Implementação

A fase de Desenvolvimento iniciou-se, não com dados de teste, como é comum, mas sim com dados que já se encontravam em Produção.

Esta fase desdobrou-se em diversas subtarefas:

- Implementação de Métricas, filtros e *prompts* no MicroStrategy;
- Criação de Relatórios Administrativos de Gestão no MicroStrategy;
- Implementação em PL/SQL;
- Criação de algumas Views;
- Processo contínuo de análise de informação;
- Produção de documentação sob a forma de:
 - Sob a forma de Listagens, resultantes de relatórios criados;
 - Documentação técnica e para o cliente.

A maior parte da implementação traduziu-se na construção de relatórios administrativos. Estes relatórios são caracterizados por resultarem da análise dos dados numa perspectiva de gestão de informação, num formato organizado em modo gráfico.

A definição de métricas e filtros no MicroStrategy permitiu aplicar as regras anteriormente descritas, de modo a que pudessem ser usadas ao nível dos relatórios. Por exemplo, estabelecer as regras temporais. As *prompts* são também para ser utilizadas ao nível dos relatórios, permitindo aos utilizadores escolher intervalos de visualização de resultados mostrados pelos relatórios. Deste modo, os relatórios fazem uso destas definições e “por trás” é gerado código em PL/SQL. Esta geração de código é uma mais-valia dado que pode ser gerado outro código em PL/SQL de teste aos relatórios que posteriormente poderá ser incluído ao nível dos relatórios. Para implementação do código criado em PL/SQL foi usada a ferramenta TOAD.

De modo a enquadrar a implementação concretizada descrevem-se as linguagens de programação SQL e PL/SQL utilizadas neste trabalho.

4.3.2.1. SQL: Structured Query Language

O SQL é uma linguagem que fornece uma interface para sistemas de bases de dados. Esta foi desenvolvida pela IBM nos anos 70 e é uma linguagem norma-padrão. Nesta linguagem estão definidos um *Data Definition Language* (DDL), usado para criar e modificar tabelas e outras estruturas da base dados, como por exemplo views e sequências, um *Data Manipulation Language* (DML), usado para efectuar a gestão dos dados e um *Data Control Language* (DCL) destinado a controlo de dados.

Em seguida apresentam-se exemplos de **DDL** usados:

- **ALTER**: alterar uma estrutura de objectos na base de dados;
- **COMMENT**: adicionar comentários ao dicionário de dados;
- **CREATE**: criar objectos na BD;
- **DROP**: apagar objectos na BD;
- **GRANT**: configurar privilégios de utilizadores às BD;
- **REVOKE**: remover privilégios de acesso definidos com o comando **GRANT**;
- **TRUNCATE**: remover todos os registos de uma tabela, sem apagar as suas definições.

Neste projecto foram usadas, com maior frequência, expressões **DML**, dado que se adaptam à análise e gestão da informação:

- **SELECT**: visualizar dados de registos existentes na BD, em conjunto com **DISTINCT** retira dados repetidos;
- **CALL**: chamar um procedimento PL/SQL;
- **DELETE**: apagar os dados de uma tabela, mas o espaço mantém-se;
- **EXPLAIN PLAN**: explicação do caminho de acesso aos dados para compreensão e optimização de código;
- **LOCK TABLE**: controlo de concorrência;
- **INSERT**: inserir dados numa tabela;
- **UPDATE**: actualizar dados dentro de uma tabela;

Foram também usadas expressões condicionais que espelham a definição de regras previamente definidas:

- **WHERE**: inserir expressões condicionais
- **LIKE**: encontrar semelhanças (por ex. palavras começadas por 'L%');
- **BETWEEN**: encontrar intervalos de valores
- **GROUP BY**: agrupar por características comuns
- **JOIN**: fazer intersecções de valores, pode ser **INNER**, **FULL**, **LEFT** ou **RIGHT** join.

Como exemplo de **DCL** tem-se:

- **COMMIT**: guardar alterações efectuadas;
- **ROLLBACK**: restaurar a BD original antes do último **COMMIT**;

4.3.2.2. PL/SQL: Procedural Language/Structured Query Language

O PL/SQL é uma extensão procedimental à linguagem SQL da Oracle. Esta linguagem suporta a existência de *arrays*, variáveis, condições, cursores e excepções. Permite a criação de uma interface com a BD de modo a que qualquer *query* SQL possa chamar um procedimento/função PL/SQL, ou então que um evento DML dispare um *trigger* PL/SQL.

Processos PL/SQL, que podem consistir em funções, procedimentos, pacotes ou *triggers*.

4.3.2.3. Análise e Levantamento de Informação

Nesta fase de Desenvolvimento foram efectuadas análises e levantamentos de informação com base na informação extraída dos relatórios administrativos de gestão construídos e decorrentes da programação em PL/SQL. Estas listagens destinaram-se à análise dos objectos obtidos como resultado da aplicação das regras inicialmente definidas. Deste modo, foi possível filtrar e identificar objectos sem uso recente, desactualizados ou obsoletos.

Foi ainda concretizada uma análise categórica dos relatórios existentes com base numa nova regra, que surgiu somente nesta fase – por periodicidade de refrescamento de informação (mensal, diária ou quinzenal).

4.3.2.4. Produção de Documentação

Para representar a informação extraída de relatórios construídos ou resultante da programação desenvolvida, foram criadas listagens de objectos e das características que derivam da estrutura do modelo dimensional. Esta documentação derivou de um processo de análise contínua da informação em questão.

O cliente é um interveniente sempre presente no âmbito deste projecto, deste modo, adoptaram-se formatos de listagens que auxiliaram os processos de análise e percepção da informação. Foram usadas folhas de cálculo Excel, *pivot tables* e *pie charts* que tornam mais fácil e intuitivo o processo de reanálise por parte do cliente. Estas listagens foram acompanhadas de documentação técnica, de modo a explicitar as regras e metodologia adoptadas, descrevendo o âmbito da informação resultante do *Reporting* administrativo construído.

4.3.2.5. Metodologia Modular

A metodologia de ciclo de vida do projecto é modular dado que inclui o cliente no processo de análise. A tomada de decisão para questões de implementação, a reanálise de listagens de resultados de *Reporting* produzidos, a validação do âmbito das regras aplicadas e a adaptar, bem como a validação de pertinência de objectos no sistema são alvo de apreciação por parte do cliente. A análise, gestão, tratamento e produção de resultados são fornecidos aos clientes para uma segunda análise.

Finda a fase de Desenvolvimento, o cliente apresenta-se como um interveniente decisivo no processo de tomada de decisão, tendo disponível um modo automático de *Reporting* administrativo, a documentação técnica e todos os levantamentos de informação produzidos. Toda a análise e documentação produzida foram enviadas para uma avaliação de pertinência de objectos e validação de regras e resultados obtidos. A documentação produzida, as *pivot tables* e modelos gráficos que acompanharam os resultados enviados garantiram que processo de reanálise, por parte do cliente, fosse intuitivo, organizado e de qualidade.

Para clientes com mais experiência na ferramenta MicroStrategy, é possível “correr” os relatórios implementados, modificando a abrangência da informação ou restringindo as regras através da aplicação das *prompts* definidas ao nível dos referidos relatórios.

Os resultados enviados ao cliente resultam de relatórios, utilizadores e tabelas cuja aplicação de regras implementadas permitiu identifica-los como não tendo utilização. Sendo, deste modo, relevantes para análise de pertinência no Repositório de Metadados. Estes objectos foram analisados pelo cliente e devolvidos com base nas regras iniciais e nas definidas pelos mesmos ao nível do Sistema de *Reporting* administrativo de gestão – definição de *prompts* e filtros ao nível dos relatórios desenvolvidos.

4.3.3. Passagem a Qualificação

A passagem a Qualificação ocorreu aquando da a recepção de feedback do cliente. Terminado o processo de reanálise, revisão das regras e abrangência dos resultados foram validadas as listagens de objectos a apagar do repositório de metadados. Num processo de *backtracking* foi feita uma revisão dos impactos e das dependências existentes, no sentido de garantir que não iriam surgir problemas decorrentes da remoção futura de objectos obsoletos.

Foi definido um período temporal de um mês para a indisponibilização de objectos para o cliente e utilizadores. Neste intervalo de tempo foi possível identificar se algum dos objectos viria a ser requisitado, confirmando a validade do processo de remoção a ocorrer na fase de produção. Manteve-se um estreito contacto com o cliente de modo a validar estas questões.

4.3.4. Testes

A fase de testes ocorreu em paralelo com a fase de passagem a Qualificação, durante a fase de indisponibilização dos objectos. Foram verificados os condicionalismos da remoção efectiva de objectos.

Os testes efectuados consistem numa avaliação diagnostica, sendo denominados de Testes de Não Regressão. Têm como finalidade a avaliação e verificação da abrangência de um sistema, ou seja, se as alterações efectuadas durante a fase de desenvolvimento não comprometem nem danificam o comportamento do sistema de Reporting.

Nesta fase foi produzida documentação técnica a manter na instituição, para um trabalho de maiores dimensões que usará todo o *Reporting* construído e as listagens de análise e documentação de objectos produzidas no decurso deste projecto.

4.3.5. Entrada em Produção e Manutenção

Na sequência dos testes, torna-se necessário efectuar uma avaliação do desempenho do sistema, depois da sua entrada em Produção. Esta entrada entra em vigor após a descontinuação de todos os objectos de Produção.

Foram removidos todos os objectos devolvidos pelo cliente como não sendo utilizados, estando desactualizados ou obsoletos no Sistema de *Reporting*.

Nesta fase final, foram efectuados cálculos estatísticos em termos de validar ganhos em memória física de armazenamento físico, que se traduz em ganhos de performance do Sistema de *Reporting* e numa diminuição da sua carga.

4.4. Detalhe dos Resultados Obtidos

Nesta secção apresenta-se o detalhe dos resultados obtidos, aplicados a um caso de estudo bem definido, no decurso deste projecto. É feita uma distinção entre resultados quantitativos e qualitativos. Apresentam-se todos os cálculos efectuados e as medidas de avaliação dos mesmos, de acordo com conceitos de trabalho relacionado.

4.4.1. Resultados Qualitativos

As listagens produzidas no âmbito de *Reporting* consistiram nos resultados com maior visibilidade do processo de exploração conceptual, análise e gestão de informação. Consistiram no elo de ligação com o processo de reanálise do cliente, numa metodologia modular. Toda a documentação técnica produzida reporta o âmbito e funcionalidades de reutilização do trabalho realizado, servindo de apoio ao trabalho futuro que faz uso deste projecto.

A documentação concretizada nas diversas fases do ciclo de vida do Projecto, consiste numa base de conhecimento que descreve uma primeira abordagem ao problema existente, numa perspectiva de análise e definição de regras num processo de gestão e identificação de objectivos. Numa fase final, a documentação produzida reflecte todo o âmbito do Projecto, servindo de base para uma futura reutilização do Sistema de *Reporting* administrativo de Gestão, após a entrada no ambiente de Produção.

4.4.2. Resultados Quantitativos

Nesta subsecção apresentam-se os dados quantificáveis ao nível de objectos do Repositório de Metadados.

Em termos de números efectivos destaca-se o tipo de objecto e a sua contabilização. Esta informação foi obtida através da construção de *queries* SQL, utilizando a ferramenta TOAD.

Objecto	Nº Total de Objectos Existentes no Repositório de Metadados
Relatórios	10 579
Utilizadores	6 912
Tabelas	6 013

Tabela 9 – Nº Total de Objectos existentes no Repositório de Metadados

O caso de estudo escolhido, no âmbito desta dissertação, foi direccionado para o tipo de objecto Relatório, uma vez que é aquele cujos resultados são mais representativos, mensuráveis de modo quantitativo e adaptáveis ao nível da avaliação dos mesmos. Para além disso, ao nível do âmbito deste Projecto, é o que mais se salienta dado que é o objecto último de restituição de informação aos utilizadores finais de um Sistema de *Reporting*, sobejamente descrito ao longo desta dissertação.

Ao nível do Repositório de Metadados da arquitectura de informação, procedeu-se à remoção de um conjunto de objectos que, segundo o cumprimento das regras de exclusão, se identificaram como sendo obsoletos.

Num total de 10 579 relatórios foram removidos os que se indicam na tabela abaixo.

Objecto	Nº Total de Objectos Removidos
Relatórios	2 884

Tabela 10 – Nº de Objectos indicados para Remoção do Repositório de Metadados

A ferramenta de *Reporting* MicroStrategy possui uma Base de Conhecimento com um conjunto de conteúdos. De entre estes, contém informações de carácter estatístico que prevêem o espaço ocupado em memória, de acordo com o tipo de objectos existentes. As fórmulas que se apresentam em seguida, foram extraídas desta Base de Conhecimento e permitiram efectuar estimativas sobre o espaço ocupado por objectos do Repositório de Metadados. Numa fase final, calculou-se o espaço ocupado após a remoção de objectos. Deste modo, obtém-se uma perspectiva comparativa de espaço ocupado inicialmente e após as remoções, comprovando o cumprimento do requisito aplicacional de redução de espaço de armazenamento físico em memória, inicialmente objectivado.

Em termos de memória ocupada a fórmula é a que se descreve em seguida:

$$\text{Memória Ocupada} = \text{Espaço Reservado} + \text{Nº Total de Objectos} * \text{Peso por Objecto Existentes}$$

- **Espaço Reservado:** corresponde a 800 KBytes
- **Peso por Objecto:**
O objecto corresponde a relatório. A Base de Conhecimento estima um peso de 93 Kbytes por relatório, correspondendo ao espaço ocupado, em média, por um relatório.

Assim, a fórmula estatística aplicada aos relatórios corresponde à seguinte:

$$\text{Memória Ocupada (Kb)} = 800 \text{ Kb} + \text{Nº Total de Objectos Existentes} * 93 \text{ Kb}$$

A contabilização do número de relatórios totais, os que foram efectivamente removidos e os resultantes após a remoção de relatórios sem utilização é a que se apresenta na seguinte tabela síntese. [Tabela 11]

Relatórios no Repositório de Metadados:	
Nº Total de Relatórios Existentes	10 579
Nº Total de Relatórios Removidos	2 884
Nº Total de Relatórios após Remoção	7 695

Tabela 11 – Relatórios Existentes no Repositório de Metadados

O número de objectos removidos do Repositório de Metadata, depois do processo de reanálise do cliente, foi de 2 884 de uma base inicialmente indicada, após análise e definição de regras, de 1 864.

Análise Comparativa: $2884 - 1864 = 1\,020$ relatórios de diferença

Esta diferença em termos de objectos indicados vs recuperados ocorreu aquando da devolução de objectos a remover pelo processo de reanálise da informação pelo cliente. A diferença quantitativa derivou de um alargamento do âmbito das regras aplicadas ao nível dos relatórios “*corridos*” pelo cliente. Uma filtragem analítica das listagens de objectos enviadas ao cliente, para avaliação da pertinência no Sistema de *Reporting*, fez com que o cliente aumentasse o intervalo temporal, definido como regra de remoção de objectos sem utilização. Daí resultaram 1 020 relatórios acessórios a remover do Repositório de Metadados.

4.4.3. Memória Física Ganha em Memória

Em seguida apresentam-se os cálculos efectivamente concretizados:

- **Memória Ocupada antes da Remoção**

$$800\text{KB} + 10\,579 * 93\text{ KB} = 984\,647,000\text{ KB} = \\ 961,569\text{ MB} = \mathbf{0,939\text{ GB}}$$

- **Memória Ocupada pelos relatórios a remover: 2 884**

$$2\,884 * 93\text{ KB} = 268\,212,000\text{ KB} = 261,926\text{ MB} = \mathbf{0,256\text{ GB}}$$

- **Ganhos em termos de armazenamento em memória após a remoção**

Após remoção o nº de relatórios corresponde a:
 $10\,579 - 2\,884 = 7\,695$ Relatórios

Em termos de espaço físico ocupado em memória:

$$800 + 7\,695 * 93 = 716\,435,000\text{ KB} = 699,644\text{ MB} = \mathbf{0,683\text{ GB}}$$

O que corresponde a 72,76% do espaço inicialmente ocupado antes da remoção, traduzindo-se numa percentagem de redução de espaço de armazenamento de **27,24%**.

Em suma, decorrente desta análise, salientam-se os seguintes resultados quantitativos:

Caracterização do Requisito Aplicacional: Diminuição do Espaço Físico em Memória	
Memória Ocupada antes de Remoção	961,569 MB = 0,939 GB
Memória Ocupada pelos relatórios a remover	261,926 MB = 0,256 GB
Requisito aplicacional atingido: Ganho em Memória Física de Armazenamento	699,644 MB = 0,683 GB

Tabela 12 – Requisito Aplicacional de Diminuição da Carga de Armazenamento

Resumindo, foram removidos 2 884 relatórios, validados numa análise conjunta com o cliente, de um total de 10 579, ficando 7 695 relatórios, com uso frequente, distribuídos por todos os Projectos das áreas financeiras do Banco.

Num Sistema de *Reporting* típico especificam-se os tempos de espera máximos admissíveis para se devolverem os relatórios aos utilizadores finais. Estes relatórios reflectem pesquisas e actualizações efectuadas ao nível da arquitectura da informação. A sobrecarga em termos de armazenamento físico é um dos factores responsáveis para o aumento dos tempos de espera.

A eficiência do armazenamento é medida com base no número de bytes a armazenar por relatórios, em memória física. Deste modo, o requisito aplicacional de redução de carga de armazenamento físico em memória revela-se como sendo da maior importância, tendo sido atingido através da redução de 261,926 MB de espaço utilizado, traduzindo-se numa redução efectiva de 27,24% do espaço em memória inicialmente ocupado.

4.5. Precisão e Abrangência dos Resultados Obtidos

Depois da aplicação da metodologia, é necessária uma avaliação das estratégias aplicadas ao caso de estudo descrito anteriormente. Um Sistema de *Reporting* aproxima-se muito de um Sistema de Restituição de Informação (Information Retrieval), onde o objectivo é devolver informação aos clientes/utilizadores finais. Para os utilizadores este processo de entrega de informação é completamente transparente, no entanto, nestes sistemas é comum contabilizar o número de documentos devolvidos e efectuar análises de avaliação de resultados obtidos.

Neste Projecto, a entrega de informação sob a forma de documentos/relatórios pode ser avaliada com base nos critérios de eficiência de execução e de armazenamento, eficácia de resposta e outras relacionadas com os clientes, procurando garantias de um sistema cujo ambiente de *Reporting* tem qualidade e um bom desempenho.

Com base em trabalho relacionado [25], neste trabalho mediu-se a eficácia/performance do sistema com base nas medidas de precisão e abrangência, vulgarmente usadas no âmbito dos Sistemas de Restituição de Informação. Fez-se uma adaptação às fórmulas existentes para cada uma das medidas, de acordo com as necessidades do Projecto.

A abrangência é uma medida de eficácia que corresponde à fracção de documentos/relatórios relevantes no total dos devolvidos. Foram efectuados testes informais de não-regressão com a finalidade de avaliar e verificar a abrangência do sistema.

A precisão é uma medida de eficácia que corresponde à fracção de relatórios relevantes, existentes nas listagens produzidas, que foram devolvidos.

Tanto a precisão como a abrangência tomam valores entre 0 e 1. Como normalmente se avalia um sistema com base nestes dois parâmetros, é comum a utilização de gráficos Abrangência – Precisão. Nesses gráficos constata-se que estas duas medidas variam numa proporção inversa, ie, quando a precisão aumenta a abrangência diminui e vice-versa.

Os resultados avaliados obtidos centram-se em duas vertentes:

- Vertente de análise realizada nas primeiras fases da metodologia;
- Vertente de validação dos resultados obtidos, avaliação dos anteriores e dos devolvidos pelo cliente, decorrentes do processo de análise de pertinência de objectos, de acordo com uma metodologia modular.

4.5.1. Cálculos da Precisão e Abrangência:

Os cálculos de precisão e abrangência foram efectuados com base nas fórmulas anteriormente descritas, adaptadas às características do Projecto. Para a categoria de relatórios relevantes, consideraram-se aqueles que foram indicados para reanálise do cliente e que assentam nas regras inicialmente definidas. Os relatórios resultantes da reanálise do cliente são os que se categorizaram como recuperados.

Em seguida apresentam-se os cálculos efectuados, adaptados ao trabalho desenvolvido ao longo do tempo.

A precisão e a abrangência, típicas, definem-se como resultando da aplicação das seguintes fórmulas:

$$\text{Precisão} = | \text{Recuperados} \cap \text{Relevantes} | / | \text{Recuperados} |$$
$$\text{Abrangência} = | \text{Recuperados} \cap \text{Relevantes} | / | \text{Relevantes} |$$

Em que:

Recuperados (relatórios resultantes da reanálise do cliente)

Relevantes (relatórios resultantes da análise inicial)

Em termos efectivos, o número de relatórios categorizados como recuperados e relevantes são representados de acordo com o projecto da área financeira a que pertencem, numa lógica de agrupamento de acordo com a ferramenta de *Reporting*. [Tabela 13]

Projecto	Recuperados	Relevantes	\cap
Projecto 1	1	1	1
Projecto 2	1	1	1
Projecto 3	7	7	7
Projecto 4	7	7	7
Projecto 5	15	16	15
Projecto 6	24	29	24
Projecto 7	41	41	41
Projecto 8	68	0	0
Projecto 9	358	135	135
Projecto 10	534	786	500
Projecto 11	1 828	841	838
Total	2 884	1 864	

Tabela 13 – Resumo de Resultado por Projecto

A aplicação imediata das fórmulas típicas de precisão e abrangência é a que se demonstra em seguida.

Precisão	Cálculos	Abrangência	Cálculos
1	1/1	1	1/1
1	1/1	1	1/1
1	7/7	1	7/7
1	7/7	1	7/7
1	15/15	0.9375	15/16
1	24/24	0.8276	24/29
1	41/41	1	41/41
0	0/68	0	0
0.3771	135/358	1	135/135
0.9363	500/534	0.6361	500/786
0.4584	838/841	0.9964	838/841

Tabela 14 – Aplicação directa das fórmulas de Precisão e Abrangência

Numa visualização em gráfico de precisão vs abrangência, tem-se os seguintes resultados. [Figura 20]

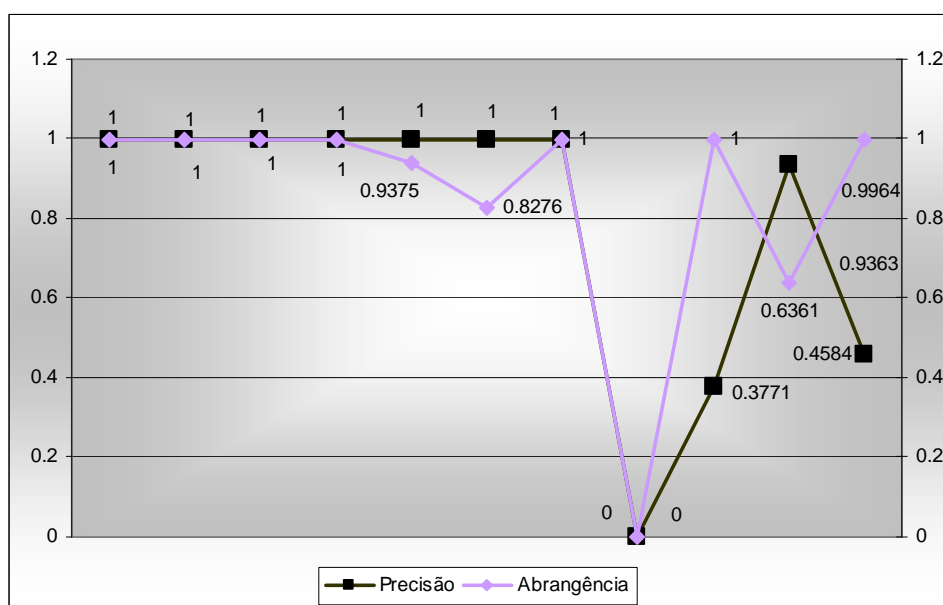


Figura 20 – Gráfico de Precisão e Abrangência dos Resultados obtidos

O gráfico demonstra que a aplicação directa das fórmulas não evidencia de modo claro a variação em proporcionalidade inversa das medidas de precisão e abrangência.

A adopção de um ajustamento às fórmulas base demonstra a variação da precisão e abrangência. O facto de ter dados com pouca variação entre relatórios relevantes e recuperados indica que a aplicação das regras foi concretizada com sucesso, não existindo uma grande diferença entre os relatórios indicados para remoção e os que foram recuperados após a revalidação do cliente. Contudo, existe uma diferença de 1 020 relatórios que derivaram de uma filtragem de resultados, pelo cliente, mais fina do que aquela que se tinha definido anteriormente, sendo uma alteração aos requisitos iniciais.

O ajustamento aos cálculos efectuados traduz-se nos seguintes resultados. [Tabela 15]

Precisão = $\frac{ \text{Recuperados} \cap \text{Relevantes} }{ \text{Somatório Recuperados} }$
Abrangência = $\frac{ \text{Recuperados} \cap \text{Relevantes} }{ \text{Total Relevantes} }$

Em que:

Somatório Recuperados: $\sum_{i=1}^{1361} \text{Recuperados}$
Total Relevantes (total absoluto)

A aplicação das fórmulas de precisão e abrangência, como adaptação à fórmula original é a que se demonstra em seguida. [Tabela 15]

\cap	Σ Recuperados	Precisão	Cálculo Precisão
1	1	1	1/1
1	2	0.5000	1/2
7	9	0.7778	7/9
7	16	0.4375	7/16
15	31	0.4839	15/31
24	55	0.4364	24/55
41	96	0.4271	41/96
0	96	0.0000	0/96
135	231	0.5844	135/231
500	731	0.6840	500/731
630	1361	0.4629	630/1361

\cap	Total Relevantes	Abrangência	Cálculo Abrangência
1	1864	0.0005	1/1864
1	1864	0.0005	1/1864
7	1864	0.0038	7/1864
7	1864	0.0038	1/1864
15	1864	0.0080	15/1864
24	1864	0.0129	24/1864
41	1864	0.0220	41/1864
0	1864	0.0000	0/1864
135	1864	0.0724	135/1864
500	1864	0.2682	500/1864
630	1864	0.3380	630/1864

Tabela 15 – Aplicação das fórmulas de Precisão e Abrangência com ajustamento

Numa visualização em gráfico de precisão vs abrangência, tem-se os seguintes resultados. [Figura 21]

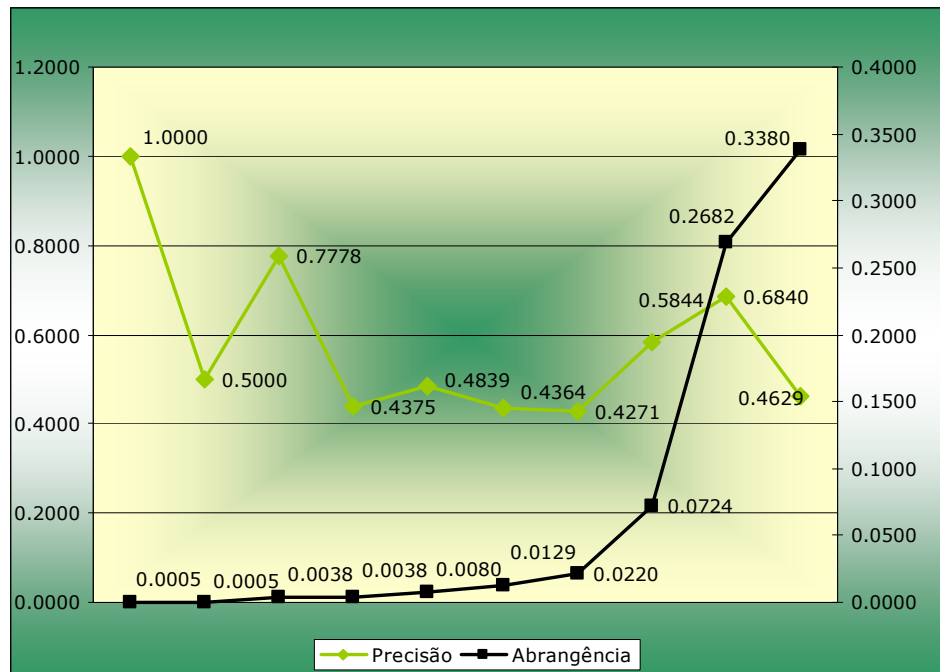


Figura 21 – Gráfico de Precisão e Abrangência dos Resultados obtidos fórmulas adaptadas

Os resultados do gráfico demonstram que a precisão e a abrangência evoluem de modo a que se a abrangência aumenta, a precisão diminui e vice-versa. A precisão e a abrangência variam, assim, em proporção inversa espelhando no gráfico a variação comum destas duas medidas de avaliação da eficiência dos resultados obtidos neste trabalho.

4.6. Principais Contribuições

De modo a sumariar as contribuições que foram sendo mencionadas no decurso deste documento salientam-se as que assumem maior destaque:

- Produção de um sistema automático de análise e gestão de informação presente num Repositório de Metadados, através da concretização de um ambiente de *Reporting* administrativo de gestão;
- Análise e identificação de um conjunto de objectos desactualizados ou sem utilização actual, através do estabelecimento de regras precisas;
- Possibilidade de reutilização do ambiente de *Reporting* de gestão administrativa concretizado. A concretização deste trabalho garante um modo automático de gerir a informação existente, a identificando e solucionando problemas ao nível da desactualização e acumulação de informação ao longo do tempo;
- Avaliação dos resultados obtidos com base nas medidas de precisão e abrangência dos mesmos, o que permitiu comprovar a eficácia do Sistema de *Reporting*, resultante da aplicação da metodologia seguida neste projecto;
- Remoção efectiva de 2 884 objectos sem utilização;
- Redução de 27% do espaço de armazenamento físico inicialmente ocupado;
O que se traduz numa redução da carga ao nível do espaço físico de armazenamento em uso;
- Integração deste trabalho num de maiores dimensões a realizar de futuro, no âmbito de migração de objectos para novas plataformas arquitecturais e nova ferramenta de *Reporting*.

4.6. Análise Crítica

Nesta secção faz-se uma análise crítica às ferramentas e tecnologias utilizadas. O trabalho concretizado e os resultados obtidos são também alvo de uma apreciação crítica.

A ferramenta de Reporting MicroStrategy, no processo de criação de relatórios gera código PL/SQL. Muitas vezes este código foi enriquecido com outro construído na ferramenta TOAD. Muitas vezes somos, assim, confrontados com a necessidade de contornar algumas limitações das ferramentas. Por exemplo, se uma ferramenta não tem uma utilização inicial muito intuitiva somos sempre levados a considerar se não seria mais rentável programar directamente em código PL/SQL, para utilizadores com experiência nesta linguagem. No entanto, com o decurso do trabalho as ferramentas tornam-se uma parte integrante do quotidiano de trabalho.

O uso continuado de ferramentas e tecnologias permite identificar potencialidades e limitações das mesmas.

4.6.1. Ferramentas e Tecnologias

Depois de se ter descrito as ferramentas e tecnologias usadas no âmbito deste projecto, em seguida, apresenta-se a análise crítica das suas funcionalidades, identificando pontos positivos ou negativos. [Tabela 13]

Funcionalidade: Reporting		
Tecnologia/ Ferramenta	Pontos Positivos	Pontos Negativos
MicroStrategy	<ul style="list-style-type: none"> ▪ Ferramenta destinada a <i>Reporting</i>. ▪ Automatização do processo de restituição de informação através de reporting. ▪ Permite uso via Web; ▪ Permite uso via Desktop para query ad-hoc. ▪ Fácil integração de informação via Reporting; ▪ Definição de <i>prompts</i> que permitem ao utilizador escolher intervalos de valores; ▪ Funcionalidades de exportação para PDF ou Excel; ▪ Boa documentação técnica; ▪ Possui uma Base de Conhecimento (Knowledge Base) com uma quantidade considerável de informação, nomeadamente estatística. 	<ul style="list-style-type: none"> ▪ De difícil utilização para utilizadores sem experiência; ▪ Actualizações nos relatórios não são, por vezes, imediatamente visíveis. ▪ Alguns erros não estão bem documentados.
Enterprise Guide	<ul style="list-style-type: none"> ▪ Ferramenta incluída no MicroStrategy destinada a Reporting administrativo de gestão; ▪ Permite automatizar Reporting administrativo; ▪ Boa documentação técnica. 	<ul style="list-style-type: none"> ▪ Alguns erros não são bem documentados.

Funcionalidade: Processamento e Análise de Informação de Reporting		
Excel	<ul style="list-style-type: none"> ▪ Adaptada à análise, filtragem e comparação de dados; ▪ Existe muita documentação acerca do uso desta ferramenta; ▪ De utilização muito disseminada; ▪ Funcionalidades complexas são possíveis sem recurso a programação complexa; ▪ Ferramenta apoiada por uma linguagem de alto nível – VBA; 	<ul style="list-style-type: none"> ▪ Limitações ao nível de linhas (só permite 65.000 registos por <i>worksheet</i>); ▪ Nº limitado de <i>worksheets</i>; ▪ Adapta-se como uma ferramenta básica de <i>Reporting</i> mas não foi inicialmente pensada para tal, motivo que faz com que esteja limitada em comparação com o MicroStrategy.

Funcionalidade: Análise e Gestão de Informação		
PL/SQL	<ul style="list-style-type: none"> ▪ Permite uma vasta gama de funcionalidades mas carece de bastante conhecimento adaptando-se a utilizadores já com experiência de programação no âmbito de bases de dados; ▪ Existem tutoriais muito completos sobre o tema, sendo uma mais-valia. ▪ Tecnologia muito disseminada e utilizada 	<ul style="list-style-type: none"> ▪ Mensagens de erro pouco explícitas; ▪ A tarefa de <i>debugging</i> é crítica; ▪ Alguma tendência a erros.
TOAD	<ul style="list-style-type: none"> ▪ Para utilizadores experientes revela-se intuitiva e de fácil utilização; ▪ Interface gráfica sem recurso a ambiente de programação em shell; ▪ Possibilidade de exportação dos dados em formato Excel, .TXT ou com delimitação de <i>tags</i> (XML). 	<ul style="list-style-type: none"> ▪ Necessita de utilizadores experientes em programação em PL/SQL. ▪ Pouca documentação acerca do uso da ferramenta.

Tabela 16 – Ferramentas e Tecnologias – Análise Crítica

Pela análise crítica constata-se que existem sempre pontos mais ou menos vantajosos na utilização de ferramentas. Contudo, numa perspectiva global, os pontos positivos superam os menos favoráveis.

4.6.2. Trabalho Realizado e Resultados Obtidos

Como críticas mais relevantes, aponta-se o facto de se ter de recorrer sistematicamente a programação como forma de contornar limitações da ferramenta de *Reporting*. Alguma da informação a filtrar encontrava-se dispersa em diferentes tabelas sem correlação no modelo dimensional, sem prontos de agregação o que obrigou a mecanização da produção e submissão de *queries* PL/SQL. A filtragem por periodicidade de refrescamento de relatórios só foi conseguida através deste método, recorrendo-se a um processo repetitivo e não automático de filtragem de informação

Conseguiu-se um modo automático de detecção de objectos sem utilização, desactualizados ou obsoletos, através do Sistema de *Reporting* de Gestão construído, o que se revelou numa mais-valia futura, num projecto a realizar e que usará como ponto de partida todo este *Reporting* administrativo de gestão de objectos desenvolvido.

A avaliação de resultados quantitativos e qualitativos permitiu comprovar a eficácia do trabalho concretizado, usando para tal uma medida efectiva de ganhos em memória física e redução de carga do Sistema, bem como uma perspectiva de carácter analítico de abrangência e precisão de resultados. A aplicação directa das fórmulas de cálculo da precisão e abrangência revelaram-se pouco explicativas quanto à variação em proporcionalidade inversas das mesmas. Para tal, foi necessário recorrer a uma adaptação das fórmulas típicas, a uma abordagem que se aproxima do processo de recepção de resultados, na metodologia modular que se manteve com o cliente. O facto de se envolver o

factor humano no processo de reanálise dos resultados de relatórios, em termos de persistência de objectos no repositório de metadados, introduziu alguma entropia, no sentido em que houve uma redefinição das regras de filtragem de informação obsoleta, comparativamente às inicialmente definidas em conjunto com o cliente.

Sumário

Neste capítulo apresentou-se o detalhe de todo o trabalho concretizado. Foi feita uma descrição de todas as ferramentas, tecnologias e técnicas de avaliação de resultados utilizadas. Fez-se uma descrição pormenorizada do faseamento metodológico adoptado durante o ciclo de vida do projecto. Foi feita uma avaliação qualitativa e quantitativa dos resultados obtidos. As medidas de eficiência de precisão e abrangência dos mesmos tiveram ênfase neste capítulo. Finalmente apresentam-se as contribuições deste trabalho e uma análise crítica das ferramentas, tecnologias, trabalho e resultados obtidos.

Capítulo 5

Conclusão e Trabalho Futuro

5.1. Conclusão

O trabalho realizado consistiu num refinamento do ambiente de *Reporting* existente num repositório de Metadados, num ambiente DW. A motivação deste trabalho surgiu a partir da necessidade de analisar e gerir toda a informação existente, na tentativa de obter um ambiente mais organizado, com base na qualidade e eficiência do serviço de entrega de informação de negócio aos mais diversos utilizadores/clientes, sob a forma de relatórios. A redução de espaço de armazenamento físico era também uma meta a alcançar.

A presente tese descreveu toda a arquitectura e metodologia associadas a um DW e, conseqüentemente, ao sistema de *Reporting* existente, enquadrando o âmbito e áreas de abrangência deste trabalho.

A concretização deste projecto originou resultados que, depois de avaliados, comprovam que a aplicação da metodologia seguida permitiu alcançar os objectivos inicialmente delineados para este projecto, num caso típico de Engenharia.

Pessoalmente, este trabalho foi uma mais-valia a nível individual tanto a nível de experiência técnica e funcional como em toda a integração na vida profissional activa.

Sumariando as contribuições de maior relevância, decorrentes da concretização de todo o trabalho subjacente a este Projecto, apresenta-se uma listagem de todos os factores que participaram para o sucesso deste Projecto.

- Produção de um sistema automático de análise e gestão de informação presente num Repositório de Metadados, através da concretização de um ambiente de *Reporting* administrativo de gestão;
- Análise e identificação de um conjunto de objectos desactualizados ou sem utilização actual, através do estabelecimento de regras precisas;
- Possibilidade de reutilização do ambiente de *Reporting* de gestão administrativa concretizado. A concretização deste trabalho garante um modo automático de gerir a informação existente, a identificando e solucionando problemas ao nível da desactualização e acumulação de informação ao longo do tempo;
- Avaliação dos resultados obtidos com base nas medidas de precisão e abrangência dos mesmos, o que permitiu comprovar a eficácia do Sistema de *Reporting*, resultante da aplicação da metodologia seguida neste projecto;
 - Remoção efectiva de 2 884 objectos sem utilização;
 - Redução de 27% do espaço de armazenamento físico inicialmente ocupado;

O que se traduz numa redução da carga ao nível do espaço físico de armazenamento em uso;

- Integração deste trabalho num de maiores dimensões a realizar de futuro, no âmbito de migração de objectos para novas plataformas arquitecturais e nova ferramenta de *Reporting*.

Todas as contribuições mencionadas garantem a satisfação dos objectivos identificados para a resolução dos problemas inicialmente existentes. O cumprimento de uma metodologia adequada à natureza do projecto, onde se adoptaram conceitos dos Modelos Organizacional, em Cascata e do ambiente típico de Produção, garantiu a satisfação dos requisitos conceptuais, processuais, arquitectural e aplicativos. Assim, este processo típico de Engenharia evidencia o sucesso de todo o trabalho realizado.

5.2. Trabalho Futuro

Todo este trabalho de análise, implementação e gestão de conhecimento, permitiu a remoção de um conjunto de informação obsoleta ou sem utilização actual. Este processo de tratamento e pós processamento de informação de gestão prepara o Sistema de *Reporting* para um trabalho futuro de migração da totalidade de informação para uma nova ferramenta de *Reporting* e consequentes novas infra-estruturas técnica e arquitectural.

Este trabalho serve de ponto de partida para um processo de descontinuação de objectos da ferramenta MicroStrategy. Actualmente já se encontra a decorrer a primeira fase de análise e avaliação da migração para uma nova ferramenta de *Reporting* que patenteia não só os relatórios como todos os processos ETL e uma profunda reestruturação das arquitecturas física e lógica de dados. O *Reporting* administrativo de gestão de objectos, construído neste trabalho, permite a detecção de objectos sem utilização desde um determinado período temporal, expurgando automaticamente objectos sem utilização que, assim, não serão migrados para a nova infra-estrutura técnica.

Referências

- [1] *"Data Warehousing, Data Modelling & Design"* – Tutorial de MicroStrategy, MicroStrategy Incorporated, 2000.
- [2] *"Apresentação Data Warehouse"* – Apresentação do Departamento, Setembro 2005.
- [3] João Gama. *"Data Warehousing"* – Apresentação, Departamento de Ciência de Computadores, Faculdade de Ciências da Universidade do Porto, 2001.
- [4] Bill Inmon. *"The Operational Data Store"* In InfoDB, Fevereiro 1995.
- [5] Ralph Kimball. *"Fact Tables and Dimension Tables"* In Intelligent Enterprise Magazine, Março 2003.
- [6] Ralph Kimball. *"The Soul of the Data Warehouse"* In Intelligent Enterprise Magazine, Abril 2003.
- [7] Richard Winter and Rick Burns. *"Managing Data Warehouse Growth"* In Intelligent Enterprise Magazine, 2006.
- [8] Warren Thornthwaite. *"Building and Delivering BI Reports"* In Intelligent Enterprise Magazine, Abril 2006.
- [9] J. Widom. *"Research Problems in Data Warehousing"* In Proceedings of the 4th Int'l Conference on Information and Knowledge Management (CIKM), Novembro 1995.
- [10] Luís B. Gouveia e João Ranito. *"Sistemas de Informação de Apoio à Gestão"* In Principia – Publicações Universitárias e Científicas, ISBN 972-8589-43-3, 2004.
- [11] Michael Alexander and Bill Jelen. *"Pivot Table Data Crunching"*, ISBN 0-7897-3435-4, Junho 2005.
- [12] P. Couto. *"WorkShop Metodologias DW"* – Apresentação efectuada na Reunião Mensal do Departamento Aplicaçional de Gestão, Março de 2007.
- [13] Thomé B., Arnold S. *"Systems Engineering: Principles Of Computer-Based Systems Engineering"*, John Wiley & Sons, ISBN 0-471-93552-2, 1993.
- [14] Oskarsson Ö., Glass R. *"An ISO 9000 Approach to Building Quality Software"*, Prentice-Hall, 0-13-228925-3, 1996.
- [15] Stevens R., Brook P., Jackson K., Arnold S. *"Systems Engineering: Coping with Complexity"*, Prentice-Hall, 1998
- [16] Susan Gallas. *"Inmon Vs. Kimball"*, Article Published in DM Direct, September 1999.
- [17] Daniel L. Moody e Mark A. R. Kortink. *"From Enterprise Models to Dimensional Models: A Methodology for Data Warehouse and Data Mart Design"*
- [18] Inmon W. *"Building the Data Warehouse"*, New York: J.Wiley and Sons, 2nd Edition, 1996.
- [19] Raghu Ramakrishnan, Johannes Gehrke. *"Data Base Management Systems"*, McGraw-Hill, 3rd Edition, 2002.
- [20] Zahran, Sami Zahran. *"Software Process Improvement: Practical Guidelines for Business Success"*, Addison Wesley, 1998, ISBN n.º 020117782X.
- [21] Fábio Rilston Paim, Ana Elizabete Carvalho, Jaelson Brelaz de Castro. *"Towards a Methodology for Requirements Analysis of Data Warehouse Systems"*.
- [22] Stephen F. LeRoy, Jan Werner. *"Principles of Financial Economics"*, Março 2000.
- [23] R. Kimball, L.Reeves, Margy Ross, W.Thornthwaite. *"The Data Warehouse Lifecycle Toolkit : Expert Methods for Designing, Developing and Deploying Data Warehouses"*. John Wiley & Sons, 1998.
- [24] IBM. *"Getting Started With Data Warehouse and Business Intelligence"*, 1st Edition, 1999.
- [25] Manuel José Ferreira Monteiro, tese de Mestrado: *"Métodos Quantitativos em Gestão"*, Instituto Superior de Estudos Empresariais da Universidade do Porto.
- [26] *Knowlegde Base – MicroStrategy* – disponível online para utilizadores registados.

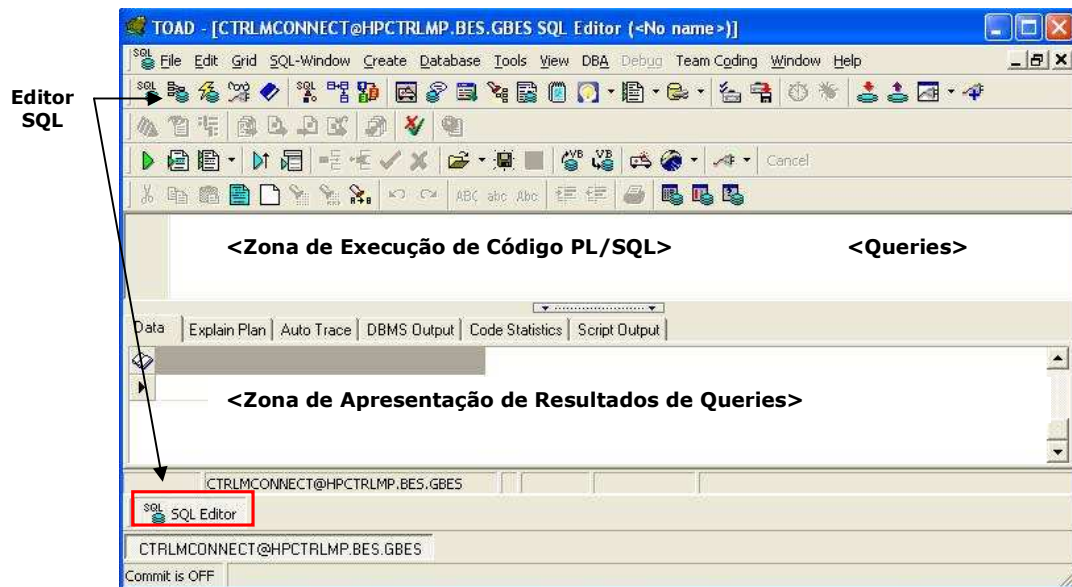
Anexos

Anexo 1 – TOAD

TOAD versão 7.6.0.11

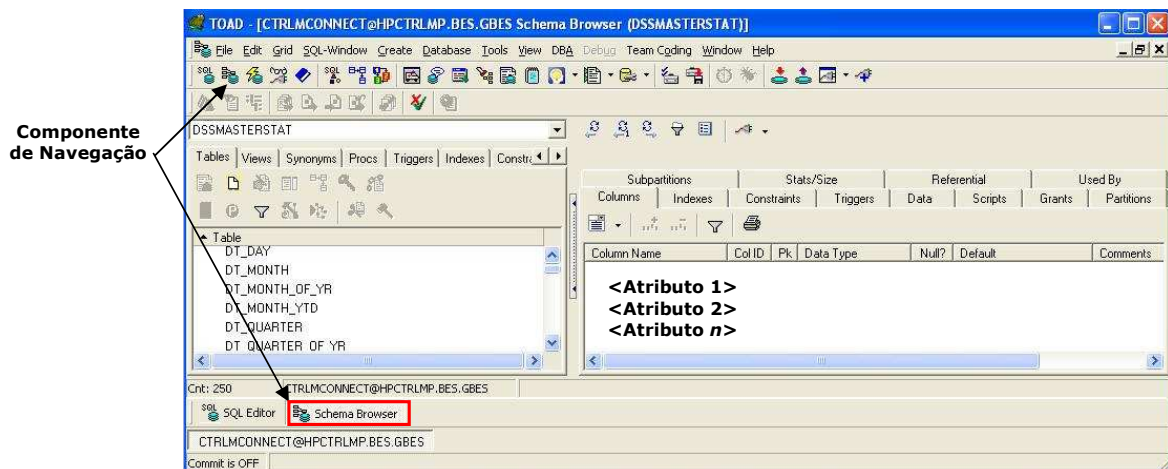
Componentes:

- Editor PL/SQL:



Anexo 1.1 - Toad: Editor PL/SQL

Componente de Navegação ou Schema Browser:



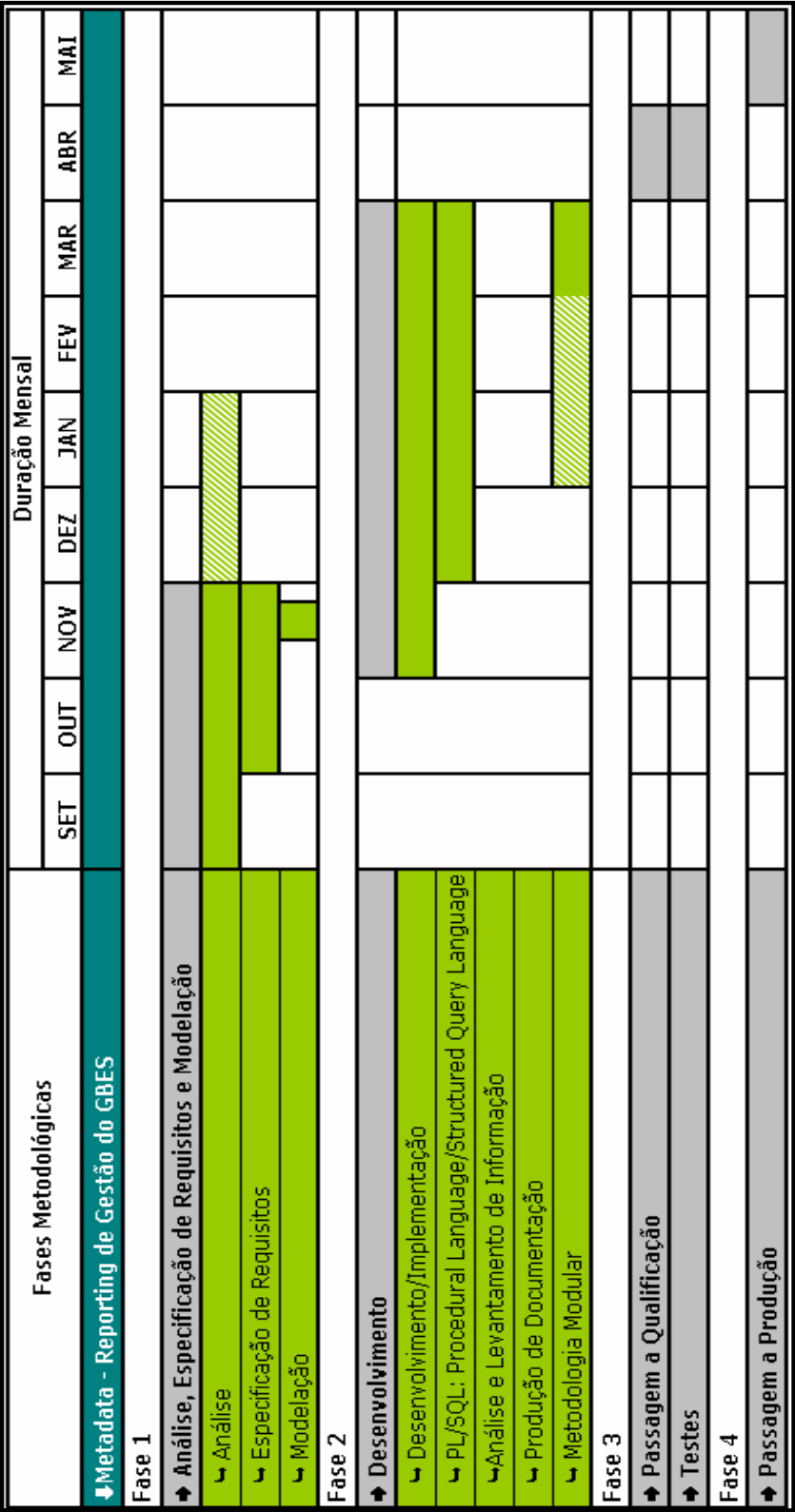
Anexo 1.2 - Toad: Componente de Navegação

Anexo 2 – Arquitectura Física

<Confidencial>

Anexo 2 – Arquitectura Física – Infra-estrutura Técnica

Anexo 3 – Planeamento: Mapa de Gantt Detalhado



Anexo 3 – Planeamento – Mapa de Gantt Detalhado